The California Central Coast Research Partnership: Building Relationships, Partnerships, and Paradigms for University-Industry Research Collaboration

FINAL REPORT
ONR GRANT NO. N00014-06-1-1111
September 30, 2006 to December 31, 2008

Principal Investigator:

Susan Opava, Ph.D.
Dean of Research and Graduate Programs
California Polytechnic State University
San Luis Obispo, CA

March 30, 2009

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
03-30-2009	Final Technical Report	9/30/06 - 12/31/08

4. TITLE AND SUBTITLE
The California Central Coast Research Partnership: Building Relationships, Partnerships and Paradigms for University-Industry Collaboration

5a. CONTRACT NUMBER

5b. GRANT NUMBER
N00014-06-1-1111

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Opava,S; Adams,N; Arakaki,D; Bensky,T; Birdsong,C; Black,M; Carpenter,T;Crockett,R; Derickson,D; Fernando,R; Gillen,K; Griffith,E; Hall,G; Harfenist,S; Immoos,C; Jansen,D; Jimenez-Flores,R; Kasper,E; Lundquist,T; Marshall,D; Meagher,J; Mehiel,E; Mello,J; Moline,M; Murray,W; Peterson,D; Rahim,A; Rosenberg,L; Schuster,P; Self,B; Sharpe,J; Slivovsky,L; Sungar,N; Szlavik,R.

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Cal Poly Corporation, Sponsored Programs Office
Bldg. 38, Rm. 102
San Luis Obispo, CA 93407-0830

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Clifford W. Anderson
Office of Naval Research
875 N.Randolph St.
Arlington, VA 22203-1995

10. SPONSOR/MONITOR'S ACRONYM(S)
ONR

11. SPONSORING/MONITORING AGENCY REPORT NUMBER

12. DISTRIBUTION AVAILABILITY STATEMENT
Available for public release

13. SUPPLEMENTARY NOTES

14. ABSTRACT
The primary purpose of this project is to carry out research and development projects and build research capacity in areas of interest to the Department of Defense and the Office of Naval Research. Research areas include nanotechnology; smart materials; communications; computing; command and control; sensors; coastal monitoring; haptics; force protection and performance; air, land and sea vehicles; neural toxins; and visual speech analysis.

15. SUBJECT TERMS
Nanotechnology; smart materials; communications; computing; command and control; sensors; coastal monitoring; haptics; force protection; human performance; ESTOL; rockets; antennae; LIDAR; neural toxins; chemical weapons; spacecraft attitude simulation; ROVs; rotors; neural networks; fuzzy logic; 3-D imaging.

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Susan C. Opava, Ph.D.
U	U	UU	UU	836	19b. TELEPONE NUMBER *(Include area code)* 805-756-1508

Table of Contents

I. Title of Project and Principal Investigator

The California Central Coast Research Partnership: Building Relationships, Partnerships and Paradigms for University-Industry Research Collaboration; Susan C. Opava, Ph.D.

II. Summary of Project

The mission of the California Central Coast Research Partnership (C^3RP) is to facilitate the exchange of technical knowledge and skills between the higher education sector and the private sector in San Luis Obispo County, and to encourage the growth of high-tech companies in the region, thereby enhancing economic development and quality of life. Since its inception, the project has focused on technologies of relevance to the Department of Defense. The partnership is a long-term plan to create a dynamic and self-supporting university-industry-government partnership that capitalizes on the strengths and mutual interests of the educational and technology-based business sectors. The plan recognizes the key role of higher education in preparing a highly skilled work force and transferring new knowledge to practical uses. The outcomes of this partnership, when fully realized, will be the creation of a robust and self-sustaining base of University R&D activities; the development of existing technology-based businesses and the creation of new ones; and the generation of opportunities for job training and research and development activities for University and Community College students and faculty.

The project also includes the construction (with non-DOD funding) of a technology park on the California Polytechnic State University campus, which will provide state-of-the-art space for private technology companies engaged in research and development activities, as well as a business incubator that will provide all of the support services needed by start-up, technology-based companies. The aspect of the program supported by this ONR grant is the continued development of a strong base of applied research at Cal Poly, through university-government-industry partnerships designed to optimize the application of the strengths of each of these sectors to problems of mutual interest. The management team, operational since January '02, continues to lead the project and develop the collaborative relationships between the educational and private sectors that are essential to realizing long-term goals and securing the financial base that will allow full-scale project development.

III. Relevance to ONR Objectives

A. Relevant partners.

C^3RP represents a coalition of educational institutions, local, state and federal government, and private businesses that have worked together in unprecedented fashion to advance the common goals inherent in the proposed university-industry partnership. The current partners in the project and their contributions include:

California Polytechnic State University

- committed the land for the construction project, valued at ~$1.5 million
- provided assistance in financial management of the project
- contributed $90,000 for a pre-feasibility study by Bechtel Corporation
- committed several hundred thousand dollars of in-kind contributions of senior management time and effort over several years and continues to do so
- invested ~$1,000,000 in efforts to raise additional funds for the project; secured sufficient private and other federal funding to construct the first building in the Cal Poly Technology Park

CENIC (Corporation for Educational Network Initiatives in California; association of Internet2 universities in CA)

- works with Cal Poly to provide high-bandwidth internet access to support C^3RP research projects

City of San Luis Obispo

- in partnership with Cal Poly has developed a carrier-neutral fiberoptic ring around the city that benefits both Cal Poly and technology-based businesses

Housing and Urban Development

- has provided funds toward construction of the pilot technology park building.

Economic Development Administration

- has provided funds toward construction of the pilot technology park building.

Efforts are ongoing to secure new partners, including:

- Major corporations
- Small technology-based businesses

B. Relevant R&D focus.

The research programs that were supported are relevant to seven of the eight "thrust areas" of ONR's Code 30 Science and Technology Program. The projects involved basic research in these areas, as well as applied research and development leading to more immediate technological applications. The seven areas of relevance and the more specific focus areas to which the research contributed are listed below:

Command and Control, Computers, Communication: situational awareness; communications; conformal antennae; visual speech analysis.

Force Protection: improved protection for the individual; smart materials; injury mitigation; IED detection; bio- and chemical-hazard detection and mitigation.

Mine Countermeasures: coastal monitoring; IED detection; situational awareness.

Human Performance, Training and Survivability: eognitive performanee enhaneement; physieal performanee enhaneement; smart materials; sensors; natural sunscreens and irradiation markers; biologieal stress reaetions; biomarkers; hapties; improved materials and processes for use on military bases and in the field.

Intelligence, Surveillance and Reconnaissance: data/information analysis and distribution; sensors; antennae.

Logistics: alternative power sourees; new materials.

Maneuver: advaneed design and materials for vehieles (land, air, and spaceeraft); ESTOL

C. University-industry-government partnership.

The primary foeus of this long-term initiative is to forge a strong link between private seetor R&D and University applied researeh to speed the development of new knowledge and the transfer of teehnology to the publie and private seetors. San Luis Obispo has beeome a draw for teehnology businesses (with a heavy eoneentration of software development companies) from both the LA Basin and Silieon Valley. For example, SRI (Stanford Researeh Institute), International operates a "software eenter of exeellenee" in the eity. Loekheed-Martin has a researeh and development group in nearby Santa Maria. Two loeal eompanies manufactured eritieal eomponents for the Mars rovers, and other companies, e.g. California Fine Wire, are suppliers to the military. Also loeated on the Central Coast are branehes of two major bioteehnology eompanies: Promega Bioseienees and Santa Cruz Bioteehnology.

D. University strengths.

Cal Poly is a State university that has aehieved national distinetion as a polyteehnie university, with engineering and eomputer seienee programs ranked among the very best undergraduate programs in the eountry. Its strengths have led it to orehestrate the researeh partnership effort and the eonsortium of partners deseribed herein. Cal Poly also has affiliations with CSA (California Space Allianee) and with Vandenberg Air Foree Base, where it has offered an M.S. in Aerospaee Engineering by distanee learning. Cal Poly has the eapability to offer many more aeademie programs by distanee learning to remote loeations. In partieular, through possible eollaborative agreements at eable-head loeations around the world (ineluding Asia and Europe) our programs ean be made available to military personnel stationed almost anywhere in the world.

Over the past several years, the University's faeulty has been turning over at a rate of about 10% a year. The University has responded to this opportunity by hiring researeh-oriented faeulty and promoting applied researeh and development. With as many as 50-60 new faeulty hires per year in the past five years, and the expeetation that this will eontinue for several more years, the University is positioned to undertake signifieant

R&D projects for government and industry. C^3RP has provided needed support and infrastructure for these faculty, which has enabled them to develop ongoing research programs and secure ~$6 of competitive funding for each $1 of C^3RP funding invested in them.

Cal Poly also has a highly qualified student body with entering credentials comparable to students who attend the highest ranked campuses of the University of California. Our students gain valuable experience working with faculty on externally sponsored research projects.

A hallmark of Cal Poly is its extensive network of industry partners. The President's Cabinet consists of more than 30 major corporate and business leaders. Each college, and each department within the college, has its own industrial advisory board. In the past, these connections were not exploited to attract industry-sponsored R&D to campus; hence, one of the goals of the C^3RP partnership is to use these existing relationships with industry to garner support for our R&D efforts, wherever possible in a three-way partnership with government entities so as to leverage the assets of each partner for the benefit of all. Cal Poly's College of Agriculture, Food and Environmental Science has successfully demonstrated this kind of partnership through its Agricultural Research Initiative. Through this initiative, a consortium of four campuses in the CSU garnered $5 million a year in on-going funding from the State of California to support agricultural research of interest to the State, with a pledge to raise matching funds from industry. This year, the State and private funding leveraged additional support from the federal government.

As will be seen in the remainder of this proposal, Cal Poly has extraordinary technical assets that can be brought to bear on the science and technology issues of import to ONR.

In summary, the California Central Coast Research Partnership has taken advantage of a confluence of factors, including existing and potential relationships, fortuitous and unique technological and economic developments in the region, the particular strengths and expertise of the Collaborative Agent Design Research Center at Cal Poly, and a meshing of the research and development interests of the University, the Office of Naval Research, and the private sector. C^3RP is the vehicle for fully realizing the benefits of the common goals and synergies of the partners and their respective resources.

IV. Summary of Results During the Period of Performance

A. General.

This program was originally funded through an award from ONR in FY '02. This report covers an award that began on 9/30/2006 and ended on 12/31/2008. General accomplishments are summarized below. Detailed reports are presented later in the document.

An overview of accomplishments during this project period follows:

- Researeh carried out by the **CADRC (Cooperative Agent Design Research Center)**, of partieular interest to ONR and the Marine Corps, was again funded. A detailed report on this projeet is provided in Seetion IV.C.1 of this report.

- **New research** has been developed and some researeh has been eontinued, ineluding some with industry eollaboration. Projeets address topies highly relevant to defense and national seeurity, sueh as nanoteehnology, imaging and analysis, smart materials, energy effieieney, eommunieations, sensors, ultraviolet irradiation, biohazards, ehemieal agents, visual speeeh analysis, artifieial neural networks, and underwater optieal sensing. Detailed reports of the results of these projeets are presented in Seetion IV.C.2 of this report.

- Sinee January 1, 2003, C^3RP-supported faeulty have **received >$37 million ($37,617,448) in competitive funding from other sources**.

- **New research capacity was developed,** ineluding new instrumentation and enhaneed infrastrueture (detailed below in Section IV.B).

- Funds were also provided to support small **student research projects** through eollaboration with Cal Poly's Honors Program. Talented, high-achieving students in Cal Poly's seleetive Honors Program were given the opportunity to work on researeh projeets with a faeulty member for 1-2 aeademie quarters and to present their results at a eampus symposium at the end of the aeademie year. C^3RP-supported projeet reports are presented in Seetion IV.C.3 of this report.

- Signifieant enhaneements have been made in **information technology infrastructure:**
    - **Internet2** eonneetivity was initially applied for, approved, and aequired for the eampus in November 2001, to support eurrent and future research efforts. Internet2 membership and eonneetivity has eontinued during this grant period.

    - We eontinued to support a **grid computing network** (IBM e-Cluster), established during a previous grant period, to support specific researeh projeets that require high-end computing eapability previously unavailable on eampus. During this period additional nodes and memory were installed. The projeets that used the environment inelude:

        1. Development and applieation of numerieal simulation teehniques to identify and optimize biomedieal device performanee.
        2. Exploring the evolutionary implieations of diatom Baeillariophyeeae) genome size variation. The primary goal of this projeet is to eompute a phylogenetie tree of speeies in the baeillariophyeeae. A maximum likelihood algorithm was used to infer the relationships based on the 18s SSU nueleie aeid

sequences. Computing power is often the limiting factor in similar projects and the ability to access a high performance system is critical to the project.
3. Exploring the effect of upstream blockage on the airflow and fuel spray in an air/fuel swirler.
4. Exploring materials or means of reducing sound and vibration of rail systems during operation.

- A **database of >70 technology–based companies** that are potential partners in the project and potential research collaborators has been updated and expanded. A series of on-campus research forums was initiated, in which these companies are invited to campus three times a year to learn about specific University research projects and identify potential areas for collaboration (see Appendix A).

- The first **research and development company to be located on campus** in anticipation of the construction of the pilot building for the technology park is flourishing in the campus environment and has developed research collaborations with faculty in several different disciplines and colleges. Applied Biotechnology, Inc. is a company that specializing in the use of genetically modified plants to produce non-food products, for example, industrial enzymes, biochemical reagents and oral vaccines. The presence of the company has spurred faculty to develop research in this area and a specialized research greenhouse was acquired to support this developing work.

- **New relationships** have been established with technology companies that are potential research collaborators, including: Aeromech, Agilight, Applied Technologies Associates (ATA), CafeFx, ConAgra Foods, Crystal Engineering Corporation, Digital West, Empirical Systems Aerospace, Epson, Form Fonts, InfoGard, Olympus, Pratt & Whitney, Propellant Solutions, Proven Management, Quintron, Rantec, Seven Pinnacles, Sonic Sensors, and Thresher Industries.

- The **web site for the project** (www.c3rp.org), which presents C^3RP as an interface between Cal Poly and business/industry for the purpose of facilitating R&D relationships, was completely redesigned and updated.

- Efforts continue to develop industry partners in the area of **alternative energy and energy efficiency** for the purpose of developing research in this field. To this end we have worked with Phycotech, AUSRA, Continental WindPower, EFuel, Energy Alternative Solutions, Inc., Pacific Gas & Electric, Rey Energy, and Blue Aqua Solutions.

- The project's leaders have continued to work with other private and government partners to advance the project and to attract research collaborators and support. During this project period, we were successful in obtaining all of the funding needed for **construction of the first building of the technology park**. Construction began in November 2008 and will be completed in Spring 2010. We also obtained letters of

intent to lease from enough technology-based companies to lease the entire available spaee (20,000 sq. ft.)

B. Development of new research capacity

One of the goals of the project was to increase the capacity of the organization to carry out state-of-the-art research in the areas of interest. To this end, specialized instrumentation was acquired and infrastructure was developed, as detailed below.

1. Instrumentation.

We acquired the following major equipment/systems. Other minor instrumentation, acquired for use on individual projects, is described in the reports for those projeets.

Research-Grade Atomic Force Microscope – This AFM was acquired on the previous ONR grant to support nanotechnology research projects. Two significant upgrades were made during the eurrent grant period to support C^3RP researeh involving the creation of nanowires and the evaluation of corrosion in microscopic materials.

Tank System - This system was acquired to allow the growth of barnacles in the laboratory for researeh on non-toxie coatings that prevent adherenee of barnaeles to marine struetures, and, therefore, biofouling. The research has applieation to materials used for ship hulls and other naval equipment and faeilities.

Electrophysiology Laboratory Equipment – An eleetrophysiology laboratory was equipped to support research on neural toxins, such as nerve gases. The equipment acquired is described in detail in the research project report entitled "Establishment of an electrophysiology laboratory for experimental validation of a combined model of the neuromuseular junction under the influence of acetylcholinesterase-inhibiting neural toxins."

Acoustic Wave and Current Sensor (AWACS) – This coastal monitoring System was aequired to support on-going marine science research at Cal Poly. It is both a eurrent profiler and a wave direetional system in one, and ean measure the eurrent speed and direction in 1-m thick layers from the bottom to the surface, as well as measure long waves, storm waves, short wind waves, or transient waves generated by local ship traffie. The system will support ongoing research including measurement of potential flow patterns for ocean contaminants, and will provide wave height/period data for use by the public. Vertical current velocity data from this sensor, combined with identieal data streams from the Cal Poly Pier, can and will be used for examining the circulation in San Luis Obispo Bay and will facilitate ONR-supported projects by E.Griffith and M. Moline. Moreover, all data will flow in real time to the NOAA regional ocean observation systems: the Southern California Coastal Ocean Observation System (SCCOOS; www.sccoos.org) and the Central and Northern Coastal Ocean Observation System (CeNCOOS; www.cencoos.org).

Real-time PCR (Polymerase Chain Reaction) System – This instrument will support several C^3RP-related genomics research projects across campus.

Laser Scanning Confocal Microscope (LSCM) – This instrument was acquired to support multiple C^3RP investigators and projects across campus. Users and applications are described below:

Lily Laiho (Biomedical Engineering)
Although well-established histological techniques are the current mainstays for the detection and diagnosis of tissue pathologies, these methodologies all require invasive excisional biopsy procedures, which invariably impose delays in diagnosis and contribute to rising healthcare costs. Thus, novel, minimally invasive, cost-effective, rapid diagnostic technologies need to be developed for disease diagnosis. Optical biopsy is one such area that is being developed to visualize and understand tissue formation, structure, and function. My overall objective is to develop an optical biopsy that enables us to study the spectral signatures of the biochemical constituents within tissue. This information can provide tissue morphological and biochemical information without the need for the removal of tissue from a patient and can be used to determine the pathological state of tissue. This system will enable sub-micron resolution that allows for imaging within intact, three-dimensional tissue and characterization of the biomarkers of normal and cancerous tissue.

John Sharpe (Physics)
In research with optical tweezers we have been trying to measure the binding forces between bacteria and proteins that occur in milk. These may provide insight into the health-promoting effects of certain components of milk. Our protocol is to coat the surface of a microscopic polystyrene bead with the protein and then, grasping the bead with the tweezers, move the bacterium and bead together until they bind. Once binding occurs, we then try pulling the bead from the bacterium. A recurrent problem we have been having is that we do not know how well the beads are coated with the protein. Is the coating uniform or does it occur in clumps? Do all beads have equal amounts of protein? Although chemical analysis can tell us that the protein is somewhere on the beads, it does not say where. With the LSCM we are able to use its excellent spatial localization properties, along with fluorescence, to see where the protein resides.

In other work we are doing related to liquid crystal devices, we have to fabricate extremely thin films (~ 2 microns) of liquid crystal between two glass plates. Variations of thickness would lead to unacceptable variations in the optical properties of the LC device. The LCSM allows us to measure thickness variations easily and accurately.

Nikki L. Adams (Biology)
The majority of research conducted in my laboratory involves examining how the environment affects development of sea urchin embryos, which are a powerful model organism. One of the main goals of our research is to assess the molecular basis of how ultraviolet radiation causes delays in cell division and development of sea urchin embryos. We have performed extensive studies examining effects of UVR on proteins

that control the cell cycle using SDS polyacrylamide gel electrophoresis and western blotting using antibodies to characterize changes in the activity of proteins. The cell cycle and early development of sea urchins is regulated by post-translational modifications of proteins including phosphorylation and subsequent localization of these proteins. Many of the cell cycle proteins are localized in the cytoplasm when they are inactive and move to the nucleus when they are active. In addition, stress-related proteins co-localize with different proteins during stressful events. We have attempted to use fluorescently labeled antibodies to identify how UVR affects localization of multiple cell cycle and stress-related proteins using a standard fluorescence microscope. Nevertheless, sea urchin eggs are yolky and vary from ~80-120 μm in diameter, preventing proper imaging of localization of these antibodies. With the confocal microscope we are able to optically section eggs and embryos of sea urchins that have been labeled with fluorescently tagged antibodies for cell cycle or stress-related proteins and identify the precise location of the phosphorylated proteins. This confocal microscope has been critical to the progression of this aspect of our research because it allows us to better identify molecular biomarkers of many types of environmental stress.

Derek Gragson (Chemistry and Biochemistry)

I have been collaborating with Rafael Jimenez in the Dairy Products Technology Center on a research project designed to develop the tools and techniques that would allow us to explore the structure of the milk fat globule membrane (MFGM). The MFGM is a natural nutrition-delivery system that has been associated with beneficial aspects of health. Over the last two decades, a great deal of knowledge on the intracellular origin, composition, and isolation procedures of MFGM has been accumulated. In recent years, several health-beneficial components of MFGM have been identified, which consequently has led to increasing interest in developing MFGM as a food ingredient with unique functional properties and health benefits. A hurdle in the path of developing MFGM as a food ingredient is the lack of basic information on the structure of this membrane derived from its lipid composition, and the interactions between the lipids and the various proteins normally accompanying it.

Our work to date has focused on exploring monolayer films derived from the phospholipid fraction of the MFGM. We have employed Langmuir-Blodgett film balance measurements, atomic force microscopy, and fluorescence microscopy in exploring these films. Our work has shown that the phospholipids phase-separate in these films, forming regions rich in phosphatidylcholine and sphingomyelin that are likely analogues to lipid rafts formed in the native MFGM. The laser scanning confocal microscope allows us to investigate this phase separation in native milk fat globules and provides us with the important link between our more fundamental, yet more easily controlled, monolayer studies and the structure in the native MFGM.

Rafael Jimenez (Dairy Products Technology Center)

In biological sciences there is an increased interest in the role that some nutrients may play in preventing or ameliorating the effect of major diseases (for example, some types of cancer, cardiovascular diseases, eye disorders, among others). In particular, our several efforts within C^3RP and ONR focus on the soldier's wellbeing. In this respect we have

focused on dairy products, and the bioavailability (the proportion of an ingested nutrient that is made available for its intended mode of action) is more relevant than the total amount present in the original food. Disruption of the natural matrix or the microstructure created during processing may influence the release, transformation, and subsequent absorption of some nutrients in the digestive tract. Therefore, the analysis of food microstructure is in the forefront of many of the projects currently being developed in my laboratory. The LCSM has enabled us to study the following: lactic acid bacteria surfaces after growth in dairy products; changes in surface of the milk fat globule membrane through production and processing of milk; survival of beneficial lactic acid bacteria after diverse drying processes in a highly nutrient-dense, ready-to-eat meal.

2. Infrastructure.

In the infrastructure area, the following projects were supported.

Research Greenhouse - New greenhouse space was acquired and equipped to enable research in transgenic plants. This new interdisciplinary area is being developed specifically for research on non-food products, in which the transgenic plant is simply the vehicle – or a protein factory - for expressing proteins of interest. The initial projects that are being undertaken focus on the production of oral vaccines in plants. These can be stored indefinitely and, when consumed, confer immunity against diseases such as malaria, traveler's diarrhea, and Hepatitis B among others. Other applications are expected to emerge, for example genetic modification of corn to greatly enhance and reduce the cost of extracting ethanol; and the production of detoxification agents (for example, against nerve gases) in transgenic plants that provide protection to the individual when the plants are consumed.

Hybrid Rocket Motor Facility – To improve safety for this C^3RP-supported project, partial funding was provided for an outdoor enclosure for the high-pressure, flammable-gas cylinders, air compressor and air dryer that are used in the experiments. These were originally located inside the laboratory.

C. Detailed research reports

The remainder of this report contains detailed individual reports of the technical results of the research projects carried out during this project period. They are presented in the following order:

1. Collaborative Agent Design Research Center (CADRC) project report
2. Other research project reports
3. Honors-student project reports

Representative publications resulting from this work are included in Appendix B.

Knowledge Management Project

Projeet Investigators:

Jens Pohl and Steve Gollery
Collaborative Agent Design Researeh Center (CADRC)
Architeeture
California Polyteehnie State University
San Luis Obispo, CA

Executive Overview

Legaey system data models ean interoperate only if their syntaetie and semantie differenees are resolved. To address this problem, researehers and students at the CADRC Center have developed the Intelligent Mapping Toolkit (IMT), whieh enables mixed-initiative mapping of meta-data and instances between relational data models.

Schema mapping eannot be entirely automated. The proeess requires human analysts to read doeumentation, to interaet with the designers and users of a system to learn their understanding of data fields, and to eoordinate with other analysts to ensure that mappings are eonsistent aeross all sehemas. IMT supports human analysis by automating speeifie aspeets using *case-based reasoning*, and enabling mappings to be defined in a eollaborative manner.

IMT employs a distributed multi-agent architeeture so that, unlike many other efforts, it ean perform mapping tasks that involve thousands of sehema elements. This architeeture includes a novel federation of matching agents that leverage case-based reasoning methods.

IMT has been tested against real world databases and has been found to be effective at reduein g both mapping time and mapping errors.

Introduction

Operational enterprise information and associated decision-support systems are components of dynamie enterprise environments. As sueh, the funetional requirements of these systems and their scope periodically ehange. New systems are introdueed to address the emerging requirements, while legaey systems undergo design modifications. This disjointed, yet organie, and ultimately evolutionary growth of eomponent systems inevitably results in data and process redundancies. Efficiencies and aeeuracies ean be gained by re-arehiteeting the eomponent systems and the assoeiated data.

As an example of this type problem, the US Transportation Command (USTC) reference data management effort is an elemental instance of the phenomenon of system redesign resulting from changes in enterprise-level requirements. Re-arehiteeting the legaey databases to eliminate redundancies and inconsistencies by adhering to the prineiple of eapturing the data only onee at the point of origin entails a laborious, error prone, and time-eonsuming task. Re-arehiteeting implies that the legaey systems may need to aeeess data and proeedures outside their native sources. Such sourees typically eonstitute the referenee data.

In the 2006-2007 projeet year, researehers and students at CADRC, under the auspiees of the California Central Coast Researeh Partnership, have partieipated in the development of the Intelligent Mapping Toolkit as a way of addressing the issues of inter-system eommunieation. IMT is intended to support analysts in eonstructing semantieally-based mappings from the data representations used by individual systems, to a common shared domain model. This report discusses the architecture and eentral concepts of the Intelligent Mapping Toolkit.

Project Description

The Intelligent Mapping Toolkit (IMT) provides a set of capacities that support analysts in mapping the data representations of legacy systems to a common shared representation. The next section describes the mapping tasks that IMT can assist with, followed by a section discussing the use-cases that the system currently addresses.

Mapping Tasks

Both the re-architected and the legacy data sources in this project constitute relational databases. Modifying the legacy applications to refer to the newly architected reference databases involves the following two sub-tasks:

1. ***Meta-data mapping*:** This entails locating tables and fields in the new database that correspond to the fields and the tables in the legacy application's database and the subsequent development of mappings between the data sources. The search space of potential mappings grows as the square of the number of entities and the attributes involved. In the USTC effort, hundreds of tables and fields are involved, which makes this a significantly complex and effort-intensive task. Manual mapping, unaided by decision-support tools, can tax resources and prove extremely difficult. Clearly, methodologies to automate mapping can improve and accelerate database interoperability efforts. However, simple mapping techniques (such as the pairing of identical field names from different databases) are bound to be of limited use because field names from tables in two independently developed databases are seldom identical.

   Variations in the naming of fields and tables can be characterized at three levels:

   - *Variations in spelling and abbreviation.*
   - *Differences in terminology arising from conceptual variations.*
   - *Ambiguous use of naming terminology.*

2. ***Instance data mapping*:** This entails the mapping of database records and fields (i.e., content) from one database to another. In many cases, the data are represented identically in both databases. However, this may not hold true when different encoding rules have been established. This typically occurs in reference data where an enlargement of scope can result in a change of encoding rules. An example of such a change is the merging of telephone number area codes or a change in telephone number format (e.g., storing phone numbers in a single phone number field versus distinct area code and phone number fields).

Current Capabilities and Design

The IMT proof-of-concept system consists of two distinct software elements: a rich client application to interface with end-users; and, a back-end object server supported by active software agents. Figure 1 provides a block-level overview of the system architecture.

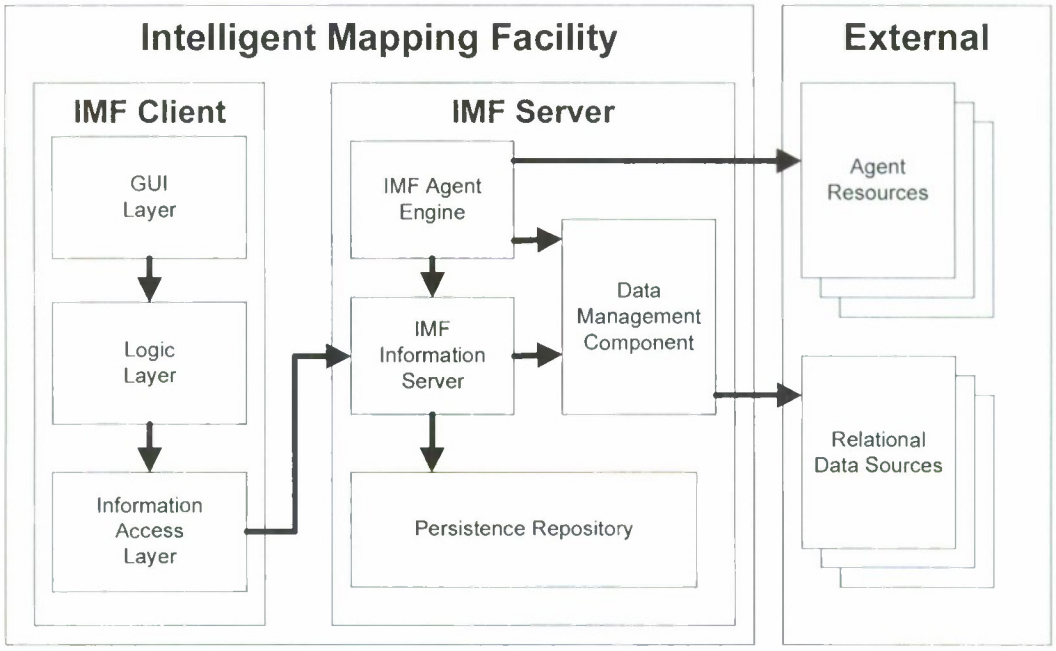


Figure 1: IMT architecture

The IMT Client is composed of three layers. The Graphical User Interface (GUI) layer provides interaction with users. The Logic layer provides the underlying application logic upon which the interface elements are based. The Information Access layer implements the interface to the information objects that the IMT Client application consumes as inputs and produces as outputs.

The IMT Server consists of an Information Server, an Agent Engine, a Persistence Repository, and a Data Management component. The Information Server provides an executable version of the domain ontology generated from its specification and accessible via a layered service-oriented interface. The domain ontology models the mapping problem inputs and an evolving solution.

High-level user interfaces conveniently encapsulate client-side access to the IMT server with other object management features such as local caching and event notification mechanisms. The available high-level interfaces for the IMT Information server ensure a large degree of physical interface compatibility with a diverse range of client implementations.

The Agent Engine interacts with the Information Server in a loosely coupled fashion that allows for opportunistic change-driven communication with the information model of the evolving mapping problem solution. An agent framework is provided that allows for a blackboard style of interaction with the information model such that multiple agents may simultaneously collaborate to produce the evolving solution representation. Agents may be easily added, removed, or substituted to allow for experimentation and scalable evolutionary system growth. Agents may utilize external agent resources such as WordNet, a publicly available linguistic ontology.

The data management component implements an abstract data access interface that utilizes the Java standard JDBC interface supported by most relational database

management systems to access the meta-data and data elements within the external relational data sources. It also provides for intelligent management of external data such that data is only imported from the external sources and loaded into the Information Server for analysis by human users and software agents as required, thereby improving overall system performance.

The external relational data sources contain the meta-data and instance data that serve as the inputs to specific mapping problems. Meta-data elements include the data model schemas associated with client systems and the USTC Master Model, representations of which are available in the system tables of SQL standard relational database management systems (RDBMS). Data elements are represented by the records within the operational levels of an RDBMS. Examples include the reference data values stored with the USTC reference data management system TMDS.

Technical Approach

For the interactive meta-data mapping task, the current system entails application of a *case-based classification methodology* to address the mapping issues that arise from spelling variations, terminological discrepancies, and ambiguous terminology usage. Classification is the task of assigning one or more predefined labels to a previously unlabelled object.

Our methodology arises from our view of the meta-data mapping task as a classification task. We view the field and tables in one database as a set of predefined labels to be assigned to the fields and tables (i.e., objects to be classified) in another database. Classification techniques inherently concern objects that share, to varying degrees, a common set of features. Many well-established classification techniques exist in the fields of machine learning, a branch of artificial intelligence, and statistical pattern recognition (Michie *et al.*, 1994). Examples include Bayesian, neural networks, support vector machines, decision-tree induction, and nearest-neighbor or case-based classification approaches.

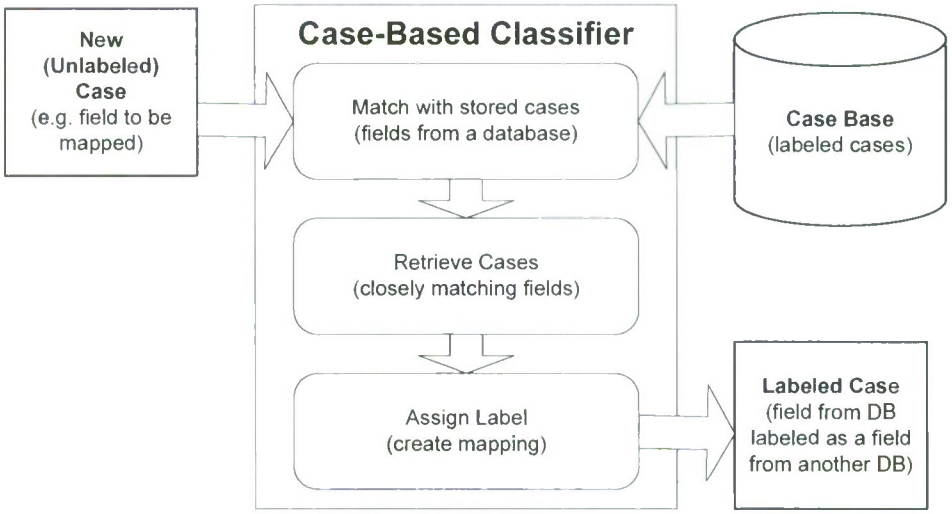


Figure 2: Case-based classification

The applicability and classification performances of these techniques depend on the representation language of objects and the amount of available pre-classified data. For example, neural networks and support vector machines work well with many examples of numerical data. In contrast, case-based techniques are well-suited for a mixture of data types with relatively few examples. The meta-data mapping task is relatively unique in that it involves numerous classes—one for each field—and provides only one example per class. Thus, only case-based classification approaches can be applied to the meta-data mapping task, in the manner detailed below.

Case-based classification works as follows. For a new object or a case to be labeled, it retrieves the most closely matching previously labeled cases from a database of cases, called a *case base*, and assigns the label from the retrieved cases as the label for the new object (Kibler and Aha 1988). For the mapping task, we define a case as the contextual attributes of a field from one database together with its name. The objective of case-based mapping is to identify closely matching fields and their context in the target database. We address the following two design factors that bear strongly on classification performance:

1. ***Case representation***: A case is a structured representation of the factors to be used for assessing similarity between two cases. The most popular case representation is a list of attributes and values. We explore various types of attributes to address the mapping issues raised earlier. For example, we test the following attributes associated with a database field:
    - The terms contained in the field name (e.g., a field name "Airport_code" has two terms "Airport" and "code").
    - Trigram components of the field names.
    - The terms contained in the table name.
    - The terms used to describe the contents of the field.
    - The database attributes (e.g., whether the field is a primary key or a foreign key).

- The properties of the field such as its length and type (e.g., numeric, string, and date).

Thus, our mapping case representation includes the attributes and the associated label (i.e., a field name).

2. ***Similarity metric***: A similarity metric is an aggregation function that assigns a number between 0 and 1 as a measure of similarity between two cases. A similarity value of 1 implies that the two cases are identical and a similarity value of 0 implies that they are completely distinct. Similarity metrics often use parameters such as attribute weights to improve classification performance. When a large number of features are associated with a case-base, some prove irrelevant and can reduce classification performance. While such parameters can be provided manually, CADRC plans to use automatic methods to achieve the same result. We select from several attribute weighting and feature selection methods at our disposal, among them information gain and rough set theoretic methods (Gupta et al., 2005). As part of our similarity metric development, we respond to terminological variations in spelling as follows:
   - *Automatically develop an abbreviation resource:* We use the USTC Master Model logical names that include non-abbreviated terms (e.g., Airport) and their corresponding physical names that include abbreviated terms (e.g., ARPT) to automatically create a lexicon of abbreviations for use in similarity assessment.
   - *An alphabetic order-based abbreviation detection algorithm*: The letters in an abbreviation preserve the relative ordering of the original word. For example, the letter "R" occurs after "A" in the term "AIR." This regularity may be exploited to detect abbreviations that are not available in the abbreviations lexicon.
   - *Detection of conceptual abstraction:* We detect conceptual equivalence of two distinct terms in the following two ways:
     i. *Automatically create a synonymous terms lexicon.* We develop this by exploiting the mapping that has already been performed. For example, if the field label "Ship Id" has been manually mapped to "Vessel Code" we can interactively identify that "Ship" is synonymous with "Vessel" and "code" is synonymous with "identifier". Such a resource is used directly for similarity assessment.
     ii. *Use WordNet* (Fellbaum, 1998), a publicly available linguistic ontology, to identify occurrences of terminological variations due to conceptual abstraction. For example, we exploit the hyponym relations (is-a-type-of) relation between concepts (e.g., Ship is-a-type-of Vessel) as part of our similarity assessment.

Agent Architecture

Below, we summarize an agent-based architecture for the implementation of the described case-based classification methodology. This architecture is depicted in Figure 3. The IMT agents operate in two distinct phases.

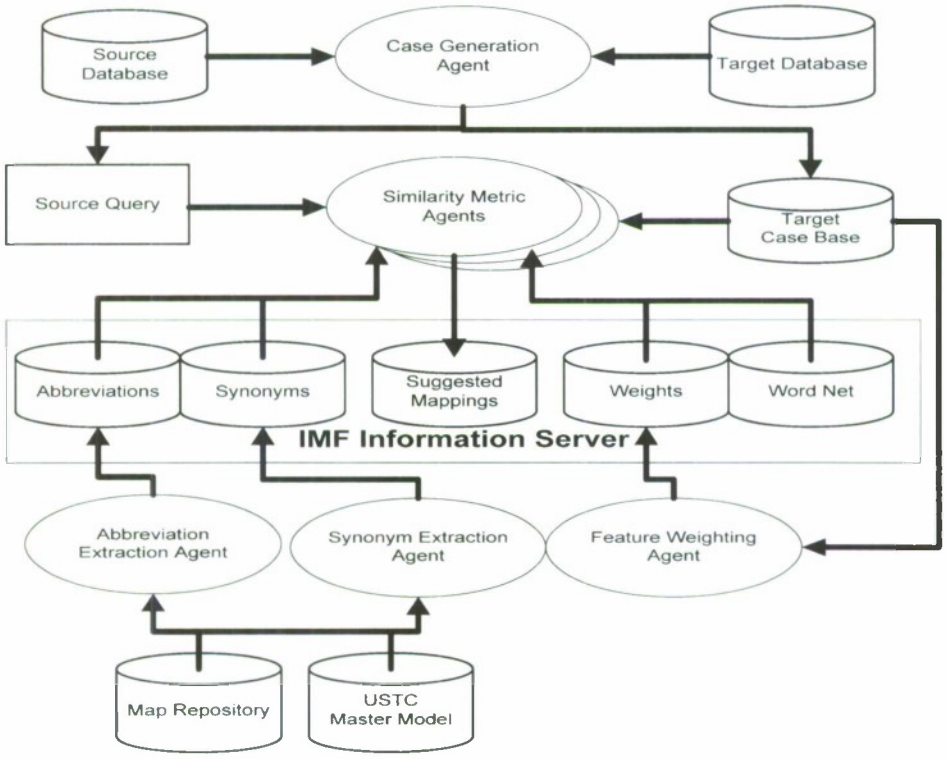


Figure 3: IMT agent architecture

1. *Resource preparation and learning phase:* In this phase, various resources for case-based classification are developed and parameter settings are learned as follows:
    a. *Case and query preparation*: The source database schema (i.e., the one to be mapped) and the target database schema (i.e., the USTC Master Model) are transformed into queries and cases respectively by the *Case Generation Agent,* which converts each field and its contextual attributes into a case.
    b. *Similarity resource preparation*: The *Abbreviations Extraction Agent* and the *Synonym Extraction Agent* use the USTC Master Model and the manual mappings respectively to develop the *Abbreviations* and *Synonym* lexicons.
    c. *Similarity metric learning*: The *Feature Weighting Agent* operates on the Target Case Base to select relevant attributes and compute their importance weights.

2. *Mapping phase:* In this phase, the *Similarity Metric Agents* treat each case from the Source as a query for semantic comparison against the target cases. These agents use the learned similarity metric along with the acquired resources to compute the similarity between the query and the target cases. The resulting scores are then presented to users as suggested mappings.

IMT Future Directions (2008)

- **Additional Agents** – The most obvious and immediate path to extend IMT capabilities is through the addition of new matching agents. Indeed IMT was designed such that new agents can be easily added without modification to the existing code-base. As discussed in the accompanying Progress Report (2 July 2007), the IMT agent framework accepts two schemas for meta-mapping problems and two table extensions for instance mapping problems. The framework performs processing on the schemas or table extensions as a whole such as the weighting of words or trigrams in regards to their rarity within the schemas or table extensions respectively comprising the meta-data or instance-data mapping problem. After this initial processing of the problem, the framework passes each pair wise schema or instance element combination to each registered agent. Agents return a similarity score for each element pair based on the agent's particular algorithmic expertise, which is typically grounded in a specific AI paradigm. The scores from the individual agents are then statistically combined to produce a composite score for each problem element pair.
    - **Context Agent** – The existing mix of IMT agents produces a high degree of meta-data match ambiguity for reference data elements. Match-result ambiguity is high when many element pairs result in near identical scores. The utilization of IMT on real-word data mapping problems indicated reasonable match ambiguity for purely operational data, but excessive ambiguity when reference data elements are introduced into the mix. Reference data is comprised of fairly static data elements intended to provide standardized enterprise-level context to system specific data. This representational division results in many specific operational data references to reference data elements. The ubiquitous use of reference data attribute names within operational data results in the pair wise combinations of the corresponding utilizations of a specific reference data attribute receiving identical – or near identical – match similarity scores. The currently implement IMT Agents utilize only table specific features when generation match similarity scores and as such do examine features capable of distinguishing the semantic differences among references to identical reference data elements. The introduction of an additional agent that examines the features of data in those tables directly referenced by table of the data element match candidates. This approach introduces the requisite distinctions

of context to greatly reduce the match ambiguity observed for reference data elements.

- **Data Value Agent** – The metadata of many domains was found to be of such poor quality that insufficient semantic context is available to the existing metadata mapping agents to provide useful mapping suggestions. Semantic features of interest for meta-data comparison include the: table and field names, free text table and field descriptions, and field data-types. If cryptically shortened 'programmer names are utilized, no descriptions provided, and most data a variable length character string there is little the agents can do to assist in the mapping regardless of the level of intelligence they employ. For this case additional features must be introduced to disambiguate the match results. The only feature that will suffice in the context of exceeding poor semantic database design and specification is the associated instance data. An agent capable of comparing the actual field values of meta-data elements is required for these cases and will additionally improve the mapping accuracy in the context of semantically rich meta-data as well.
- **Other AI paradigms** – The open-ended agent framework provides an ideal test bed for experimentation with a diverse breadth of AI paradigms (e.g., neural networks, expert systems, and descriptive logics), for their applicability to data mapping problems. The automated regression testing framework that was incorporated into IMT provides the ideal means to quantify agent performance against standard test cases and thus compare the utility of differing paradigms as applied to the generalized meta- or instance-data mapping problems or to those within specific specialized communities of interest.

- **Automated Agent Weighting Factors** - IMT allows end-users to adjust the weighting factors of individual agents to control the effects of a specific agent's similarity score results on the composite agent result. This approach recognizes that different agents – based on different approaches or paradigms – may exhibit different performance (i.e., result accuracy) characteristics in different domains. While adjustable weighting factors provide powerful domain customizability, determining appropriate weighting factors is not a straight-forward trivial exercise for typical users. This problem can be resolved by utilizing previously generated data mappings within a given domain in conjunction with a training algorithm to automatically determine a good mix of weighting factor values for the given domain.
- **Service Orientation** – The IMT was designed and implemented as a stand-alone single-user system. However, much of the underlying agent-based mapping capabilities could be useful in other application contexts. In this regard, the underlying intelligence of the semantic mapping engine should be factored out into a standalone Web service that could be accessed via the Internet by other

software applications to perform runtime mapping services in support of multiple COIs.

- **Improved Computational Resource Utilization** – The size of the mapping problems addressable by an IMT instance are bounded by the computer memory resources available. Currently the IMT generates all possible combinations of semantic elements and produces a similarity score for each combination. All scored combinations are retained until the user confirms a specific mapping. For a given source attribute there may be thousands of potential target attributes. By only retaining a fixed number of the top-most source target pairings the IMT memory footprint could be significantly reduced thereby allowing substantially larger mapping problems to be addressed. If the user does not find a suitable mapping in the top-ranked pairs retained, the remaining pairs can be regenerated – for just the single source attribute – on demand.
- **Mapping Code Generation** – Ultimately the purpose of mapping meta-data is to translate instance data from one representation to another. For this purpose transform code must be written. The semantic map output of the IMT is a conceptual specification only, but contains the essential information from which it should be possible to generate the transform code corresponding to an IMT semantic mapping; thereby realizing substantial savings on development costs when applying IMT to interoperability projects.
- **Runtime Instance Mapping** – Given the software code to transform one form of information representation to another, the existing instance extensions to the mapped schemas must still be correlated in order to synchronize the information in two distinct databases. The existing capabilities of the IMT InstanceMapper could be easily adapted to perform this function without human intervention.
- **Intelligent Query** – The same Case-Based Classification (CBC) technology employed by the IMT to rank the similarity of a given meta-data or instance-data element to those within another repository could be utilized to produce a powerful *query by similarity* capability. Traditional database technology utilizes exact match –Boolean logic – constraints to identify and retrieve an item of interest (e.g., name = Joe or length < 100). This can be further augmented somewhat using keywords as is commonly done with internet search engines (e.g., description contains 'Joe'). Regardless, standard technology requires one to in essence exactly know what one is looking for in order to identify and retrieve corresponding data. A CBC-enabled Query by description similarity paradigm would allow one to simply write a free form English text description of the desired data to be able to have something at lease close to the desired information returned, if available.
- **Data Cleansing** – An Intelligent Query could in turn be employed in data cleansing applications. For example, given a 'dirty' data record for a supply item. Suppose the data record requisitions a supply item specified by an invalid stock number – a requisite key to order the requested item. If the requisition also contains a description of the desired item it could be used as the argument to a

similarity query into the stock number database to identify the valid stock number for the desired item.

References

Chisholm, Malcolm (2001); Managing Referenec Data in Enterprise Databases; ISBN: 1-55860-697-1; San Diego, CA: Academic Press.

Gupta, K.M; Moore, P.G.; Aha, D.W; and Pal, S.K. (2005): Rough-Set Feature Selection Methods for Case-Based Categorization of Text Documents; S. K. Pal; S. Bandyopadhyay; and S. Biswas (Eds.); Lecture Notes in Computer Seience (LNCS 3776), (pp. 792-798) : Heidelberg, Germany: Springer.

Fellbaum, C. (1998); WordNet: An Eleetronie Lexical Database. MIT Press.

Kibler D. & Aha, D.W. (1988); Case-Based Classification; E. Rissland & J.A. King (Eds.), Proceedings of AAAI Workshop on Case-Based Reasoning, (pp.62-67); St. Paul, MN.

Michie, D., Spiegelhalter, D.J., and Taylor, C.C. (Eds.) (1994); Machine Learning, Neural and Statistical Classification; New York, NY: Ellis Horwood.

Pohl J. (2001); Information-Centrie Deeision-Support Systems: A Blueprint for Interoperability; Office of Naval Research, Workshop on Collaborative Decision-Support Systems, Quantico, VA, 5-7 June, 2001.

Pohl K. (2001); Perspective Filters as a Means for Interoperability Among Information-Centric Decision–Support Systems; Office of Naval Researeh, Workshop on Collaborative Decision-Support Systems, Quantico, VA, 5-7 June, 2001.

USTRANSCOM Corporate Information-Centric Environment (CICE) (2003); USTRANSCOM Architecture and Technical Integration Division Report, prepared by CDM Technologies, Inc. USTRANSCOM TCJ6-A; Scott Air Force Base, Illinois, Oetober 2003.

Zang M. (2003); The Knowledge-Level Approach to Intelligent Information System Design; InterSymp-2003, 15th International Conference on Systems Research, Informatics and Cybernetics, Baden-Baden, Germany, 29 July- 1 August, 2003.

Molecular Sensors and Defenses Against Ultraviolet Radiation

Project Investigator:

Nikki L. Adams
Biological Sciences
California Polytechnic State University
San Luis Obispo, CA

Molecular Sensors and Defenses against Ultraviolet Radiation

1. ABSTRACT

Many organisms are exposed to damaging levels of ultraviolet radiation (UVR, 290-400 nm). The overall goals of this study are to identify molecular targets of UV-induced damage and to characterize and quantify natural sunscreens (mycosporine-like amino acids, MAAs) in local marine organisms, using sea urchins as a model organism. We proposed the following: **Aim 1.** Identify whether Checkpoint kinase 1 in sea urchins, *Strongylocentrotus purpuratus* (SpChk1), which is directly activated by UVR in other organisms and which blocks mitosis, can be used as a marker of UV-induced damage in sea urchin embryos. **Aim 2.** Screen and characterize the function of natural sunscreens, mycosporine-like amino acids (MAAs) in local marine algae and sea, so we may ultimately examine whether molecular targets are protected by MAAs and whether MAAs can be developed as bio-indicators of UV-exposure and broadband sunscreens.

2. RESULTS

Aim 1-Identify molecular targets of UV-induced delays in cell division in sea urchin embryos, *Strongylocentrotus purpuratus* as a model.

Optimization of one-dimensional gel electrophoresis methods for identifying SpChk1. We examined whether Chk1 kinase is present and affected by UVR in sea urchin embryos. Chk1 is a kinase that controls activity of Cdc25 and ultimately Cdc2 or mitosis. We had previously attempted to examine whether the Chk1 protein is present in *S. purpuratus* using direct protein analysis (SDS PAGE and western blotting). We tested multiple antibodies for anti-human total Chk1 and pChk1, which did not adequately recognize Chk1 in *S. purpuratus*. Therefore, we utilized a more creative approach and this projeet expanded to cloning and sequencing Chk1 in *S. purpuratus*. We have now demonstrated that two antibodies we produced against *S. purpuratus* Chk1 (SpChk1) recognize a single protein in *S. purpuratus* lysates (Fig. 1). We have optimized our conditions to identify SpChk1 in lysates of sea urchin eggs and embryos.

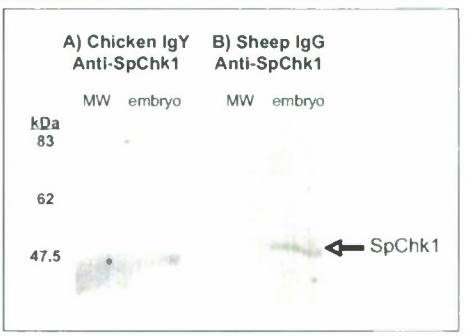


Figure 1. Immunoblots of proteins from *S. purpuratus* embryos (100 min) cell lysate (25mg/lane) (lane 1) incubated with A) 1:10000 affinity purified chicken IgY antibody to synthetic SpChk1 peptide(Aves Labs). Identified using a 1:1000 dilution of horseradish peroxidase (HRP) conjugated sheep anti-chicken (Aves Labs) or B) 1:1,000 affinity purified sheep IgG antibody to synthetic Chk1 peptide. Identified using a 1:1000 dilution of rabbit anti-sheep (MP Biomedical Inc.). Bands were detected with ECL plus (Amersham) using a Typhoon Imager (GE/Amersham). The expected size of Chk1 is 52 kDa.

Determine whether Chk1 exists in sea urchin embryos

This year, we used these antibodies and confirmed that Chk1 is available in unfertilized eggs and throughout early embryonic stages (Figure 2). In addition, we have examined whether UVR affects the concentration of total SpChk1 during early development. Embryos were irradiated using artificial UVR and assays described above. Protein lysates from various time points during the first cell division were examined using SDS PAGE and western blotting (Figure 3).

Our results indicate that SpChk1 protein is available in the egg through early development (i.e. during stages where we see UV induced delays in division) (Figure 2). In addition, we have demonstrated that UV-irradiation of eggs does not affect the concentration of SpChk1 (Figure 3)

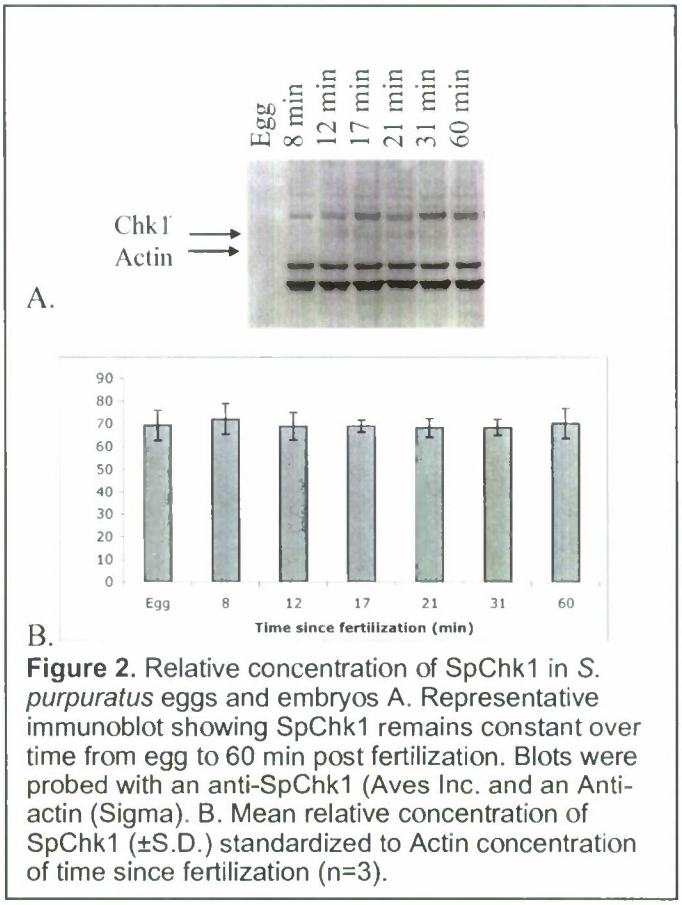


Figure 2. Relative concentration of SpChk1 in *S. purpuratus* eggs and embryos A. Representative immunoblot showing SpChk1 remains constant over time from egg to 60 min post fertilization. Blots were probed with an anti-SpChk1 (Aves Inc. and an Anti-actin (Sigma). B. Mean relative concentration of SpChk1 (±S.D.) standardized to Actin concentration of time since fertilization (n=3).

Examine whether UVR treatment affects Chk1 phosphorylation and activity in early embryonic cell cycles. We have performed many experiments examining whether UVR affects the phosphorylation state or total concentration of Chk1. Over the past year, we have attempted multiple assays to address this aim with limited success. We have attempted the following:

1. Commercially available antibodies for pChk1. We have tested ~10 commercially available antibodies against human pChk1 for sea urchin lysates. Despite our experience using phospho specific antibodies for pCdc2 that were generated against human proteins, none of the pChk1 antibodies worked reliably for our sea urchins lysates. We are not convinced that means that Chk1 is not phosphorylated as a result of UVR treatment because the most promising these antibodies would not even recognize positive controls. We tried to work with the manufacturers, but they have admitted to similar problems with these antibodies.

2. Immunoprecipitation: We planned to perform immunoprecipitations (IPs) of Chlk1 using our anti-SpChk1 antibodies along with 1-D SDS PAGE with western blotting or 2-D gel electrophoresis (2D GE) with mass spectroscopy analysis to identify the phosphorylation state of Chk1 and co-localizing proteins. We have attempted protocols provided by IP experts (Dr. K. Foltz etc.) and experienced major difficulties immunoprecipitating Chk1 using our anti-SpChk1 antibody. We have attempted to IP it using both native and denaturing conditions, but consistently get a precipitate in all samples. We have discovered that this may be due to the high PI (9.6) of Chk1 or the peptide region that our antibody was generated against and that our antibody is unlikely to work for immunoprecipitations.

3.Phosphoprotein isolation. We have attempted to use a Pro-Q® Diamond Phosphoprotein Enrichment Kit to isolate phosphorylated proteins from non-phosphorylated proteins. Our goal was to then run these isolates on a standard 1-D SDS PAGE and probe with our Chk1 antibody. Nevertheless, in using actin as a control protein (we did not expect to see changes in phosphorylation) we obtained actin in both "phosphorylated" and nonphosphorylated" fractions. We are still trying to optimize this technique. We attempted to establish a protocol to immunoprecipitate Chk1 from sea urchins protein lysates using the SpChk1 antibody produced in chickens (data not shown). Next, we will examine these immunoprecipitates using two-dimensional gel electrophoresis (2D GE) combined with mass spectroscopic analysis (MS) to characterize the phosphorylation state of Chk1 and identify co-localized proteins from sea urchin embryos. We will collaborate with Dr. Tomanek, who has successfully applied these techniques, to determine our optimal methods.

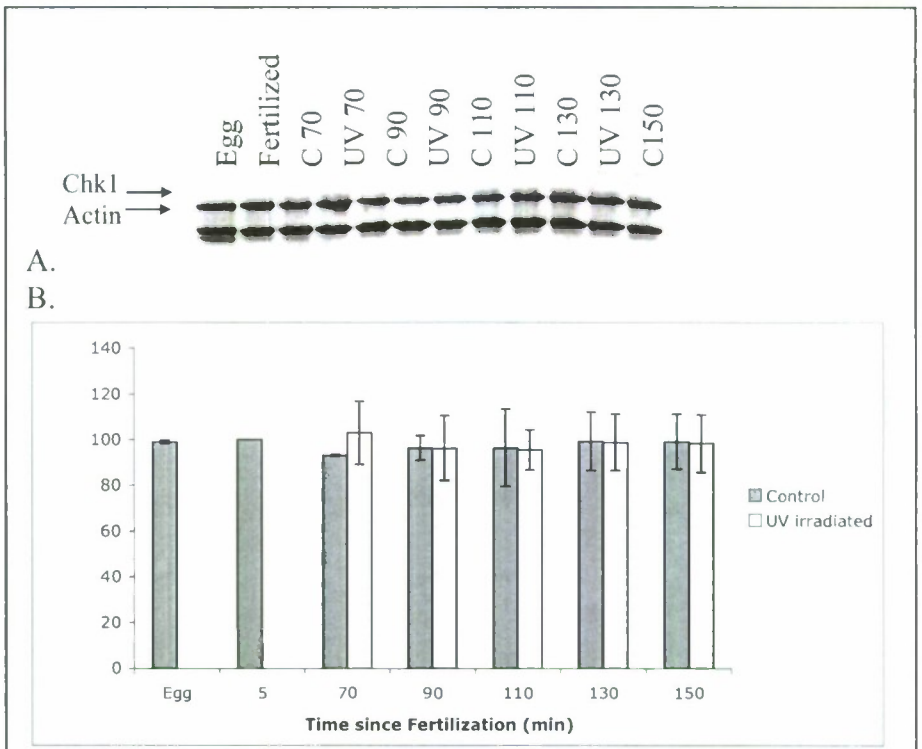


Figure 3. Relative concentration of SpChk1 in *S. purpuratus* eggs and embryos either exposed to or protected from UVR is not affected by UVR. A. Representative immunoblot showing the concentration of SpChk1 in sea urchins embryos remained constant despite UV-irradiation of embryos. Blots were probed with an anti-SpChk1 (Aves Inc. and an Anti-actin (Sigma). B. Mean relative concentration of SpChk1 (±S.D.) standardized to Actin concentration for each treatment at time since fertilization (n=3).

4. Two dimensional gel electrophoresis: We have attempted 2-D gel electrophoresis and western blotting for Chk1 on 2-d blots. Nevertheless, the high PI of Chk1 warrants using higher than standard isoelectric focusing range and a special loading chamber, which we do not have at this time. We will pursue this avenue further.

5. ***Chk1 inhibitor assays:*** We have also performed functional assays using inhibitors of Chk1 activity (GO6976 from Cal Biochem (Kohn et al., 2003)) to determine whether Chk1 regulates UV-induced delays in division in sea urchin embryos. Because Chk1 transmits the signal for a mitotic checkpoint, stalling the cell in division, we expected UV-induced delays to be lower or diminished in the presence of a Chk1 inhibitor. Briefly, we have exposed sea urchin embryos to

various combinations of UVR and concentrations of GO6976 to determine whether the inhibitors ameliorate UV-induced delays in cell division (Fig. 4, n=1). Our experiments confirmed that UVR causes delays in mitosis of sea urchin embryos. Nevertheless, aliquots of embryos incubated with Chk1 and no UVR experience a delay in division compared to non-exposed (UV or G06976), but the overall UV-induced delay in mitosis is lower than in embryos not exposed to the inhibitor. We perform further experiments with a lower concentration of GO6976. The control embryos divided at the same time, but there was no difference in the amount of cleavage delay. Therefore, this assay either tell us that Chk1 is not active in these cells, Chk1 does not affect UV-induced cleavage delays or that this inhibitor is not working in sea urchin embryos. We are exploring other Chk1 inhibitors to further explore this question.

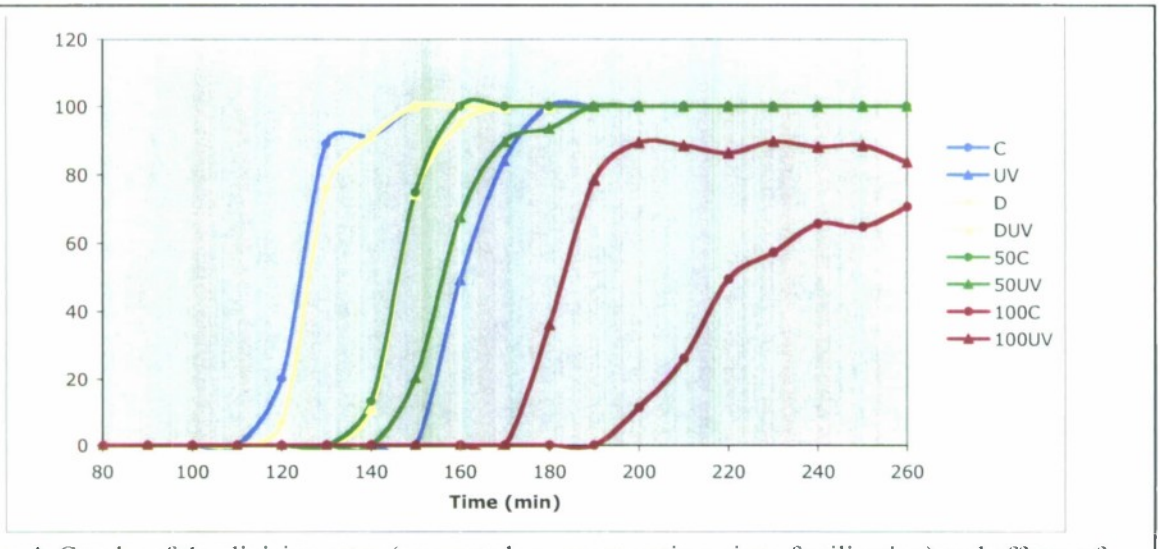


Figure 4. Graphs of the division rates (percent cleavage over time since fertilization) and effects of UVR and Chk1 inhibitor (GO6976 dissolved in DMSO) on *S. purpuratus* embryos. Treatment of eggs and embryos were as follows: C=Control (no UV, no GO6976; UV=+UV, no GO6976, D=DMSO, no UVR, DUV=DMSO, +UV, no GO6976; 50C=no UV, +50μM GO6976; 50UV=+UV, + 50μM GO6976; 100C=no UV, +100μM GO6976; 50UV=+UV, +100μM GO6976. UV-irradiation causes a delay in division in all cases. DMSO may have decreased the effects of UVR. GO6976 treatment alone appears to delay development, whereas GO6976 +UVR decreases the UV-induced delay, but may lead to decreased total cleavage.

Conclusions: Our results indicate that UVR does not affect the concentration of total SpChk1. This result is consistent with the UV-induced response in other organisms, where regulation occurs at the level of phosphorylation at Ser 296, 317 and 345. Therefore, we have attempted to examine whether phosphorylation of Chk1 is affected by UVR, but we have been unable to identify anti-pChk1 that recognizes SpChk1. We are currently attempting to use ion exchange purification of Chk1 followed with SDS PAGE and western blotting. Presumably, nonphosphorylated Chk1 will have a very high positive charge, whereas phosphorylated Chk1 will become more negative and we may be able to separate these two forms from one another using an ion exchange column. In addition, we have recently successfully expressed SpChk1-His, which will allow us to perform more functional assays on effects of UVR on SpChk1. Overall, these investigations may allow us to identify the molecular pathway between UV-induced stress and cell cycle proteins and allow us to develop molecular markers to detect UVR-induced damage in multiple marine organisms.

Aim 2. Identification and quantification of MAAs in local sea urchins and red algae to screen for novel MAAs, examine mechanisms of protection and whether MAAs may act as bio-sensors of UVR.

Previously, we demonstrated that adult *S. purpuratus* contain MAAs in their gonads, but that the concentrations are highly variable among and within collection sites. Our results indicate that some, but not all, *S. purpuratus* eggs contain potential photobiochemical defenses against ultraviolet radiation. Nevertheless, the concentration of MAAs in ovaries of sea urchins appear to vary within sites, indicating there may be some effect of microhabitat which vary from burrows, to exposed pool and south or north facing walls or proximity to specific types of algae that determines whether sea urchins acquire MAAs and whether embryos are protected against UVR damage.

During the past year a masters student conducted a more extensive investigation of whether MAAs may vary in sea urchin gonads due to microhabitat or algal content of the sea urchin diets to allow us to more effectively identify regions with sea urchins that have reliable concentrations of MAAs. She sampled a variety of tissues from sea urchins and sub-habitat (microhabitats) types at each site (e.g., exposed pools, burrowing under rocky overhang etc.) for MAAs. In addition, she is simultaneously sampling algae from each microhabitat and performing algal transects on a seasonal basis to determine the availability of particular algal types, the MAA compliments in those types of algae and whether availability and MAA concentrations vary on a seasonal basis. These studies will also allow us to identify algae to feed sea urchins controlled diets in the future to create batches of eggs with high or low concentrations of MAAs to determine whether and how MAAs protect against UV-induced molecular damage.

Briefly, she is characterizing the types and concentrations of MAAs in algal and sea urchin tissues (epidermis and gonads) from four microhabitats that correspond to microhabitats where we also find sea urchins (rocky intertidal areas that are north facing slopes, south facing slopes, horizontal surfaces where sea urchins are mobile on surfaces and horizontal surfaces where sea urchins are burrowed into the rocks(=pit sea urchins)). She collected 10 dominant algal species, 12 specimens per species at least one from three replicates of each microhabitat. Similarly, she collected sea urchins at two sample times over the past year and tested the gonads for MAAs. Tissues were excised, frozen at -80°C and then lyophilized. MAAs were extracted from equal amounts of algal samples (200 mg) using 80% methanol. Extracts were analyzed using high performance liquid chromatography as described previously (2007 report). Individual MAAs were identified and quantified using comparisons to known MAA standards prepared by Dr. Walter Dunlap at AIMS. We reported data for algal MAAs in our 2007 report. Analysis of variance and general linear models used to examine whether there were differences in MAA concentration among tissues, sexes among microhabitats and two sampling periods.

Results

Total MAA concentration per urchin tissue was significantly higher in females than in males (Figure 5a, $p<0.000$), increased significantly from November to January (Fig. 5 B, $p=0.023$), and was significantly higher in gonadal tissue than in epidermal tissue (Fig. 5C, $p<0.000$). A comparison of the changes in MAA concentration in tissues among the sexes over time shows that male tissue MAA concentrations did not change from November to January while female

tissue MAA concentrations increased (Fig. 5A) though the sex by month interaction was not significant (p=0.672). Ovary MAA concentrations were higher than testes, but epidermal MAA

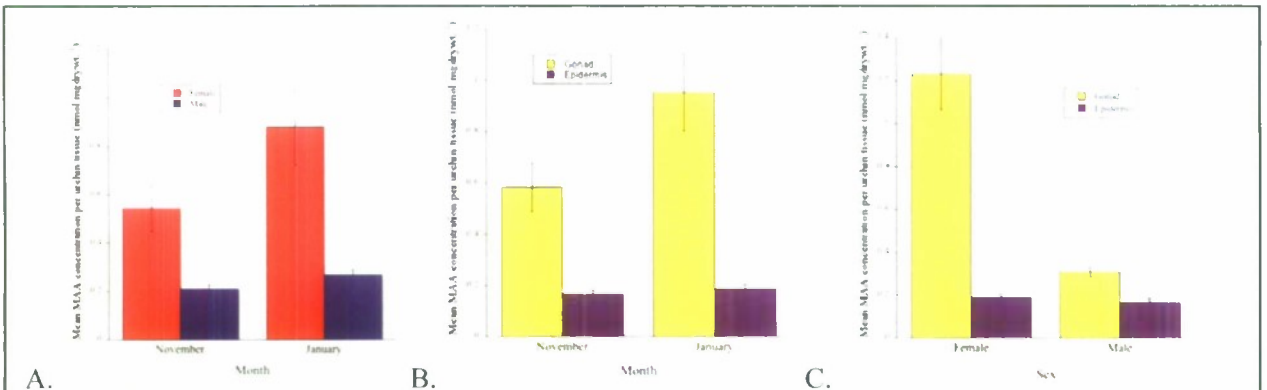


Figure 5. The mean concentration of total MAAs (± S.D.) in S. purpuratus tissues from November 2006 to January 2007. A. Total MAAs in all tissues from females vs. Males over time and B. MAAs in gonads vs. epidermis over time C. Total MAAs in gonads vs. Epidermis by sex.

concentration did not differ among the sexes (Fig 5B) though the sex by tissue interaction term was not significant (p=0.668). Comparing MAA concentrations in tissues over time shows that the gonadal concentration increased from November to January, while epidermal concentration stayed the same, though the interaction of month and tissue was not significant (p=0.074) (Fig 5C). However, when the sexes are separated, the effect of month on MAA concentration in the different tissues varies in the female sea urchins (p=0.027) and not the males (p=0.954). In female sea urchins there was a strong increase in MAA concentration in gonads but not in epidermal MAAs from November to January (Fig 5b), while in males there was little increase in gonadal MAA concentration and none in epidermal MAA concentration (Fig 5C).

There was a significant effect of microhabitat on MAA concentration in sea urchin tissues (p<0.000) with the pit sea urchins having significantly lower mean MAA concentration per tissue than the other three microhabitats, which did not significantly differ from one another (Fig 6). The effect of sex on MAA concentration in sea urchin tissues differed significantly among the microhabitats (p=0.010) and the data show that, for females, the MAA concentration per tissue drops drastically in the pit microhabitat, while in males it drops slightly (Fig 6).

The MAA concentration of sea urchins in the different microhabitats was significantly affected by tissue (p=0.044) and this relationship becomes more obvious when considering the sexes separately. In females, there is no significant effect of tissue on MAA levels among the urchins in different microhabitats (p=0.990), while in males this

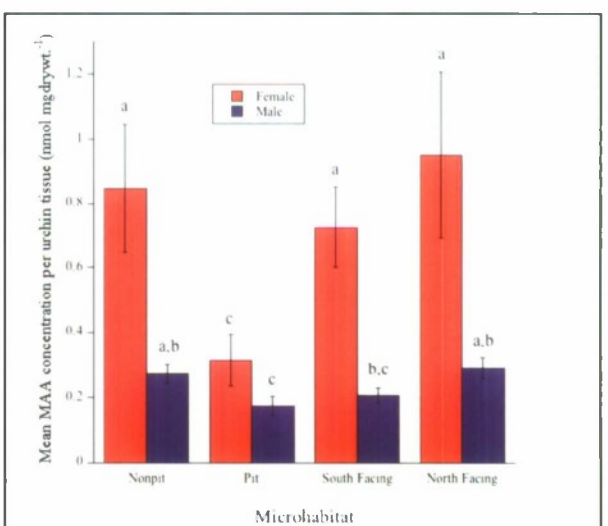


Figure 6. Mean (±S.E.) concentration of total MAAs in sea urchin tissues by sex among various microhabitats at Avila Beach Jetty, Avila Beach, CA. Individual letters indicate significant differences in MAA concentration among sample groups.

effect is highly significant ($p<0.000$). In females, mean gonadal MAA concentration is higher than epidermal concentration in every microhabitat, and both mean concentrations for both tissues are reduced in the pit urchins (Fig 7a). However, in male sea urchins the mean gonadal MAA concentration is higher than the mean epidermal MAA concentration in the shadier South facing, North facing and Pit microhabitats, but is lower in the high irradiance non-pit microhabitat (Fig 7b).

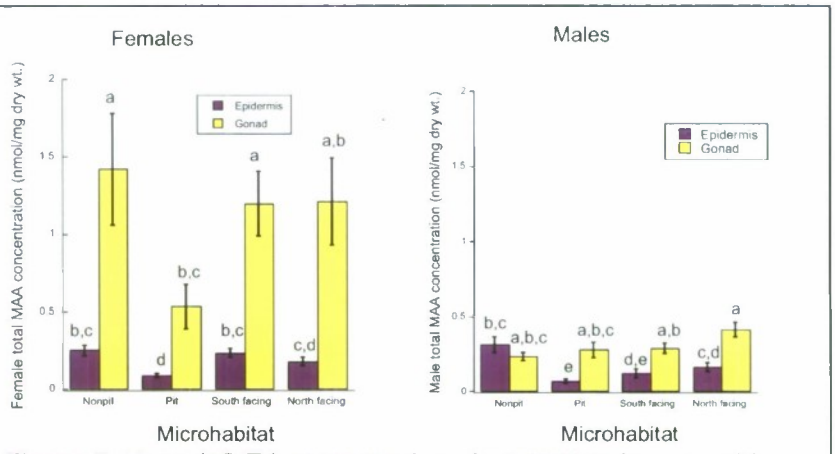


Figure 7. Mean (±S.E.) concentration of total MAAs in sea urchin tissues by sex and tissue types among various microhabitats at Avila Beach Jetty, Avila Beach, CA. Individual letters indicate significant differences in MAA concentration among sample groups.

The relative amount of individual MAAs within the tissues among the sexes was variable (Fig 8). The mean shinorine concentration (± S.E.) was highest in ovaries and in epidermal tissues for females and males (53.6 ± 28.9%, 72.8 ± 21.8% and 63.2± 28.9%, respectively) followed by palythine (38.4 ± 31.4%, 19.8 ± 22.9% and 29.5 ± 29.3%, respectively). Palythine and shinorine were the principal MAAs found in testes with palythine accounting for 88.4 ± 13.7% and shinorine only 10.8 ± 13.1%. Ovaries had a significantly higher relative amount of mycosporine-glycine than testes and epidermal tissues ($p<0.000$). Testes had a significantly higher amount of palythine than ovaries and epidermal tissues ($p<0.000$) but had significantly lower amounts of mycosporine glycine ($p<0.032$), shinorine ($p<0.000$), porphyra-334 ($p<0.000$) and asterina-330 ($p<0.000$).

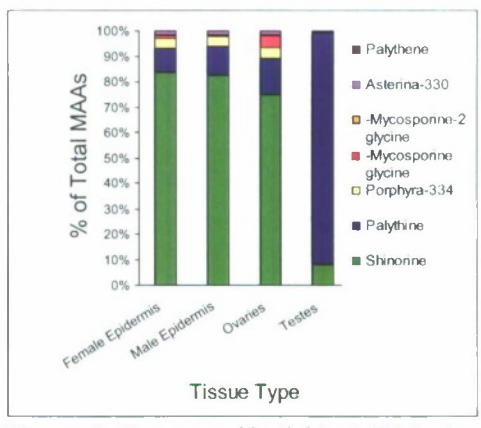


Figure 8. Percent of individual MAAs in tissues of *S. purpuratus* types varied significantly.

Conclusions

We have firmly established that *S. purpuratus* contain MAAs in their tissues. Nevertheless, the concentration can be highly variable. Our results demonstrate that microhabitat can influence that concentration of MAAs in tissues. In addition, the tissue type and sex of sea urchins also affect the concentration of MAAs. MAAs tend to be higher in ovaries than testes (as we predicted), whereas the MAAs concentration of MAAs in the epidermis does not vary by sex. Overall, concentrations of MAA are higher in gonads that epidermis, but this trend appears to change with males in some microhabitats. Sea urchins that burrow in pits tended to have lower concentrations of MAAs than those sea urchins that tend to move more freely, which could allow us to more reliably identify which urchins have high or low concentrations of MAAs in their eggs for future studies. Nevertheless, the sea urchins that are burrowed in pits tend to be smaller

overall and have lower gonadal indexes, implying that their nutritional status is weaker than other sea urchins and that such comparisons may have additional confounding factors other than just MAAs. Our results indicate that we will need to rear sea urchins on controlled diets of algae with or with out MAAs to obtain more reliable populations of embryos to examine the functional role of MAAs.

This past year, we initiated a long term study rearing sea urchins on controlled diets that are either MAA-rich or MAA-poor to identify whether MAA concentration of eggs can be enhanced or at least create a more reliable MAA concentration gradient. We plan to utilize the embryos of these populations of sea urchins with variable concentrations of MAAs to examine the role of MAAs as natural sunscreens and examine whether they protect against the molecular targets described in Aim 1.

USE OF FUNDS

Funds were used to purchase materials and supplies, pay for travel for students to two scientific meetings, provide release time for the PI, and provide assistantships for two graduate students, Sarah Striekler (Chk1) and Sarah Gravem (MAAs) and two undergraduate students. These students have presented these results at local scientific meetings in posters or oral presentations. We will present these results at further scientific meetings and publish results from this work in refereed journals with students as co-authors on these papers. In addition, these data will be used to provide a foundation for future work and grant proposals to continue this work that will be submitted to the National Science Foundation or the National Institutes of Health. Additional funding was provided to help build the aquaria for experiments rearing sea urchins on controlled algal diets at our CCMS. We appreciate this funding opportunity.

Wireless System Development of Reduced-Size Aperture Radiators, Wireless Interfaces for Biomedical Systems, and Radiated Emissions

Project Investigator:

Dean Arakaki
Electrical Engineering
California Polytechnic State University
San Luis Obispo, CA

C^{3}RP Final Report, March 2008

The primary objeetives of the 2007-2008 C^{3}RP project have been to:

- Develop HFSS eomputer models of metamaterial struetures
- Develop physieal metamaterial prototypes verified through measurements taken in the Cal Poly Aneehoie Chamber
- Develop eomputer model of human head and oral eavity for intraoral sensor simulations.
- Develop an Eleetromagnetie Compatibility (EMC) radiated emissions ehamber and establish testing eapabilities
- Charaeterize the EMC ehamber and aeeompanying instrumentation; earry out radiated emissions testing on aetual systems

Metamaterials Development

The development of HFSS simulation models for metamaterials has been established for split ring resonators (SRR) and eapaeitively loaded strips (CLS) as detailed in [1-3]. Both planar (Fig. 1)

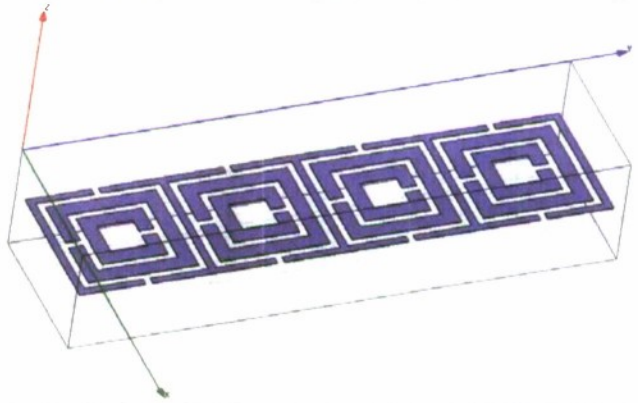


Fig. 1: Unit Cell, Planar MetaMaterial HFSS Model

and tri-planar models (Fig. 2) have been developed and simulation results eompared.

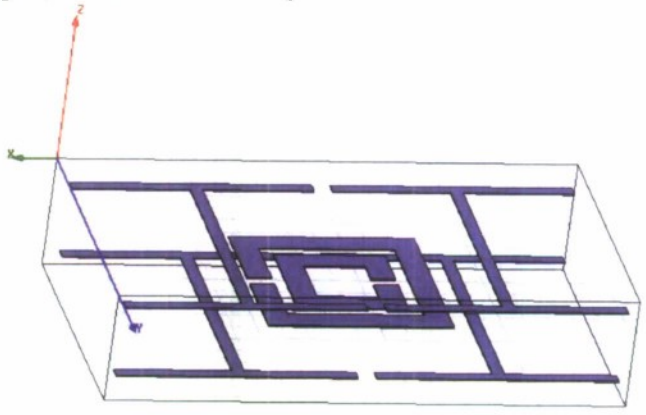


Fig. 2: Unit Cell, Tri-Planar MetaMaterial HFSS Model

The planar model (Fig. 1) is composed of a box filled with 5880 Duroid ($\varepsilon_r = 2.2$) or FR4 with (x, y, z) dimensions of (0.210, 0.756, 0.156 inches). S-parameters are calculated as a function of frequency with the goal of $|S_{11}| \rightarrow \infty$ dB and $|S_{21}| \rightarrow$ 0dB to yield a double negative material [$\mathrm{Re}(\varepsilon) < 0$, $\mathrm{Re}(\mu) < 0$] response. Fig. 3 below illustrates the S-parameter response for the planar case (Fig. 1).

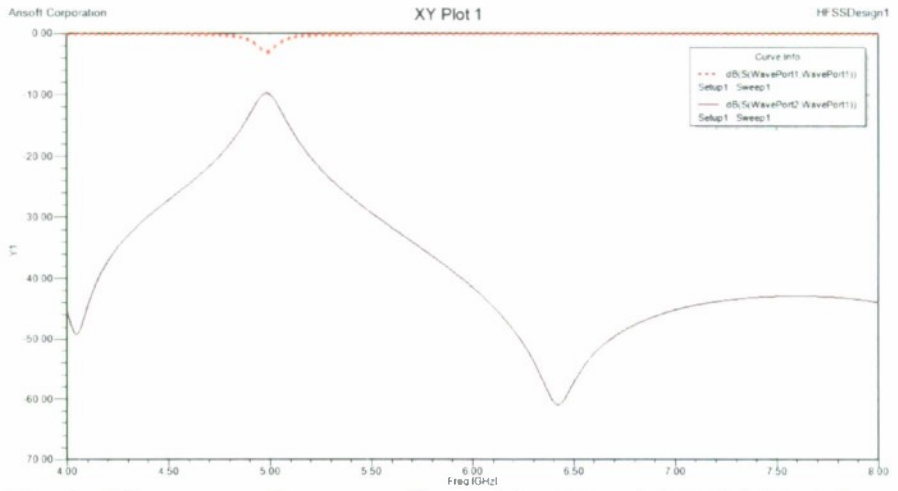


Fig. 3: S-Parameter Response, Planar MetaMaterial HFSS Model

This response indicates double negative properties ($|S_{11}|$ = -4dB and $|S_{21}|$ = -10dB) at the target frequency of 5GHz. To determine if the tri-planar structure yields an improvement in performance, the same measurement was performed: see Fig. 4 below.

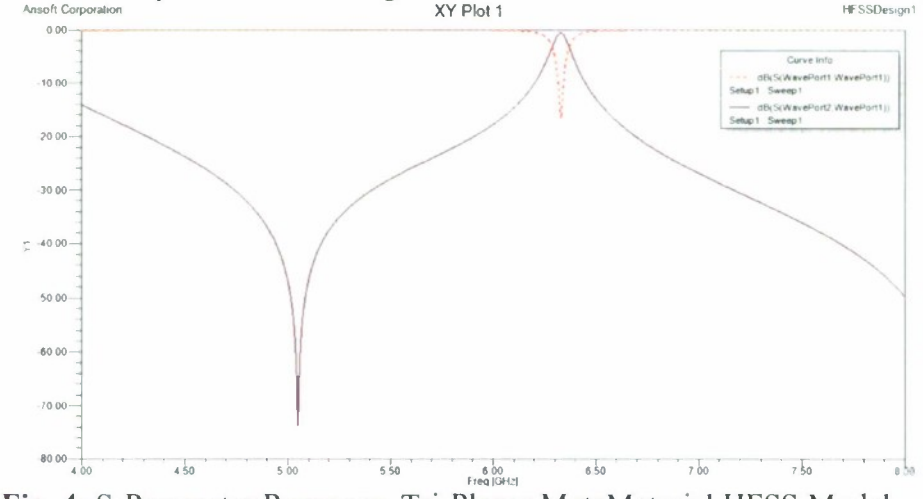


Fig. 4: S-Parameter Response, Tri-Planar MetaMaterial HFSS Model

This response also indicates double negative properties ($|S_{11}|$ = -18.5dB and $|S_{21}|$ = -0.05dB), but at a frequency of 6.325GHz. Material dimensions can be scaled to obtain these characteristics at the target frequency (5GHz). Additional simulations are currently underway to characterize negative material characteristics with respect to unit cell geometries. To allow fabrication, simulation models align with materials (Roger's Duroid and FR-4) available from substrate manufacturers.

To detect negative index material characteristics, a prism structure of metamaterial unit cells [4] will be fabricated: see Fig. 5 below.

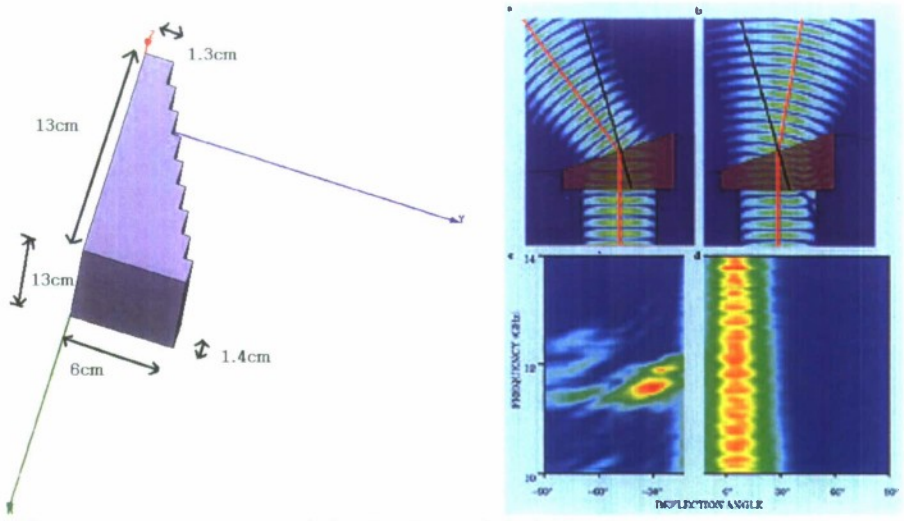


Fig. 5: Prism Geometry, MetaMaterial HFSS Model (left), Expected Ray Paths (right) [4]

Incident radiation will be applied at the x = 13cm plane at normal incidence (along the x-axis), while transmitted radiation will be measured as a function of angle (xy-plane) on both sides of the normal (at expected refracted angles) to verify negative index material performance.

Biological System (Intraoral Sensor) Development

It was discovered that the proposed intraoral sensor has already been developed by a research company and hence was abandoned in favor of the design and construction of the EMC (electromagnetic compatibility) radiated emissions chamber.

EMC (Electromagnetic Compatibility) Radiated Emissions Chamber Development

The development and construction of the EMC radiated emissions chamber [5-7] involved the following components

- extension of external copper mesh frame: 14'L × 9'W × 7.5'H to 16'L × 9'W × 9'H
- design and construction of internal 2" × 4" support structure (for ferrite tiles)
- installation of 0.5" plywood – 26AWG sheet metal – 0.5" plywood mounting surface for ferrite tiles
- ferrite tile (4" × 4" × 0.25" thick) installation: 6006 tiles
- connector panel door: allows cable entry through chamber wall
- chamber door design: tile design around main entry
- turntable and motor specification
- probe antenna specification and acquisition
- spectrum analyzer, software-controlled, specification and acquisition

To accommodate test articles with maximum dimensions 18" L × 18" W × 6" H, the chamber was extended in length by 4ft and in height by 1ft: see Fig. 6 below [8].

Fig. 6: EMC Sereen Room Expansion: Note new panel on right, 1ft height extension panels

Once the sereen room was complete, RF (radio frequeney) isolation was eharacterized to determine attenuation levels from outside to inside the newly-renovated ehamber: see Fig. 7 below.

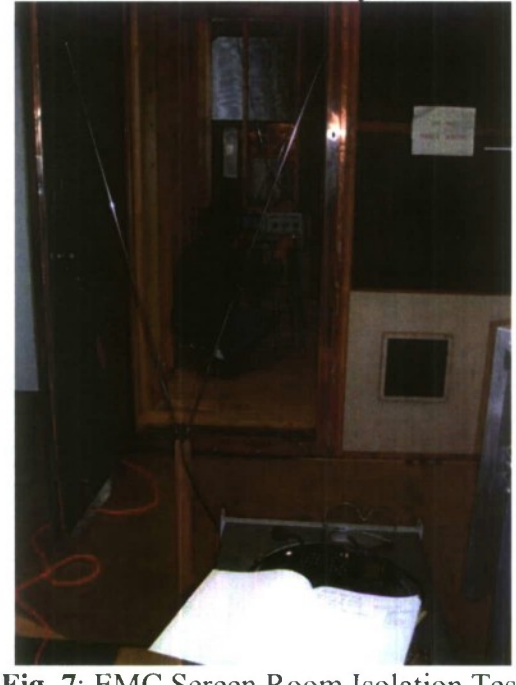


Fig. 7: EMC Sereen Room Isolation Test

The worst-ease measured isolation is 19dB at 1GHz, while the best-ease is 64dB at 100MHz. Between these frequeney limits, the average attenuation is approximately 30dB. Since commercial ehambers have minimum isolation levels of 90dB to 100dB, additional measures to inerease isolation will be employed ineluding the installation of fingerstoek to both sides of the door jamb and patehing holes in the eopper mesh. The addition of steel sheet metal between plywood layers

should improve isolation values taken without the support structure. Results will be reported in a subsequent report.

In conjunction with the screen room expansion, a wooden support structure was developed to accommodate a plywood structure (plywood - sheet metal - plywood) upon which ferrite tiles will be mounted: see Fig. 8 below.

Fig. 8: Ferrite Tile Support Structure inside EMC Screen Room

To support the plywood sandwich structure mounted on the ceiling, vertically placed 2" × 6" boards were installed across the double-stacked 2" × 4" top layer: see Fig. 9 below.

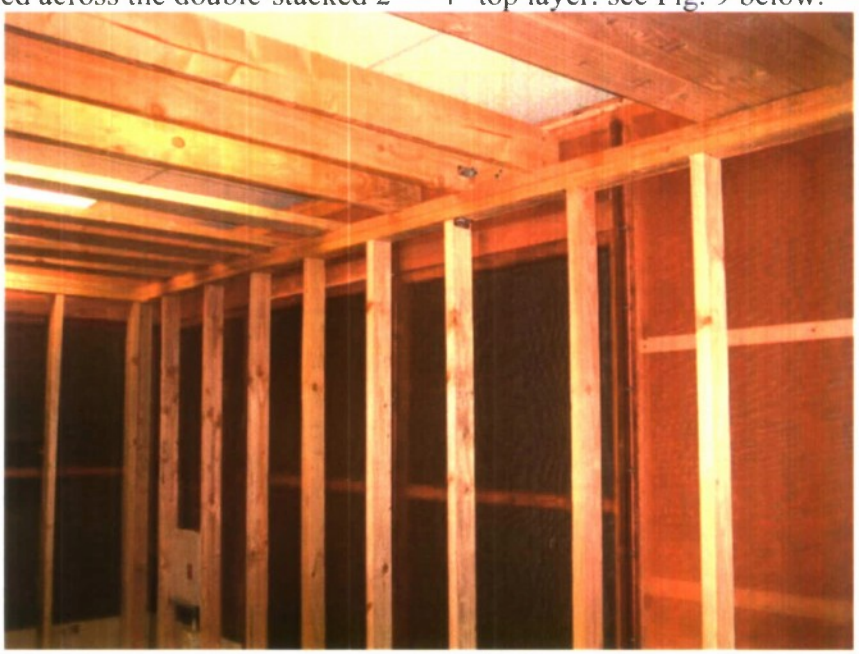


Fig. 9: Ferrite Tile Support Structure inside EMC Screen Room: 2" × 6" Ceiling Joists

To maximize absorption in the UHF frequency range, a plywood-sheet metal-plywood (½ inch thick plywood) structure was installed onto the support frame: see Fig. 10 below.

Fig. 10: Plywood – Sheet Metal – Plywood (½ inch thick) Sandwich Structure

With the mounting structure in place, ferrite tiles (4" × 4" × ¼" thick each) are attached to the structure with galvanized steel screws: see Fig. 11 below.

Fig. 11: Ferrite Tile Installation

Through C^3RP/ONR support, a total of 6,006 ferrite tiles (143 boxes, 42 tiles/box) [8, p. 107] were purchased from Samwha Corporation, in addition to 6,006 plastic inserts and mounting hardware. The chamber is currently under construction (see Fig. 12 below).

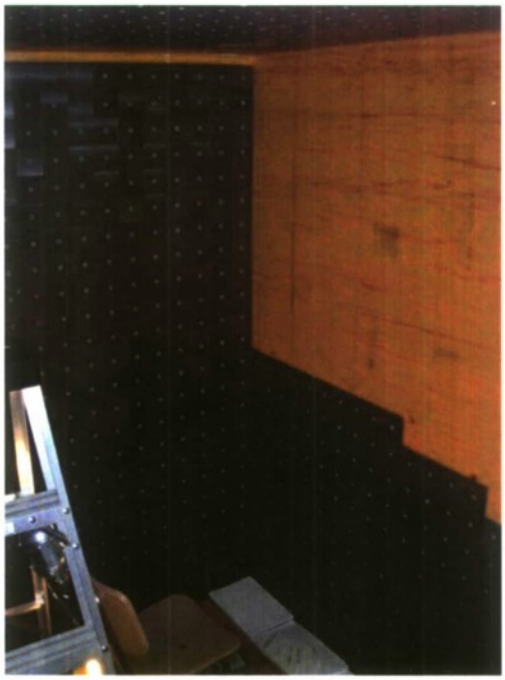


Fig. 12: Ferrite Tile Installation; Ceiling, Back, and Side Walls

A turntable, stepper motor, and angular position control system will also be designed and implemented inside the new chamber to enable radiated emissions measurements as a function of angular direction from the equipment under test (EUT). A cardboard cutout indicating the size of the proposed turntable and required location on the chamber floor is shown in Fig. 13 below.

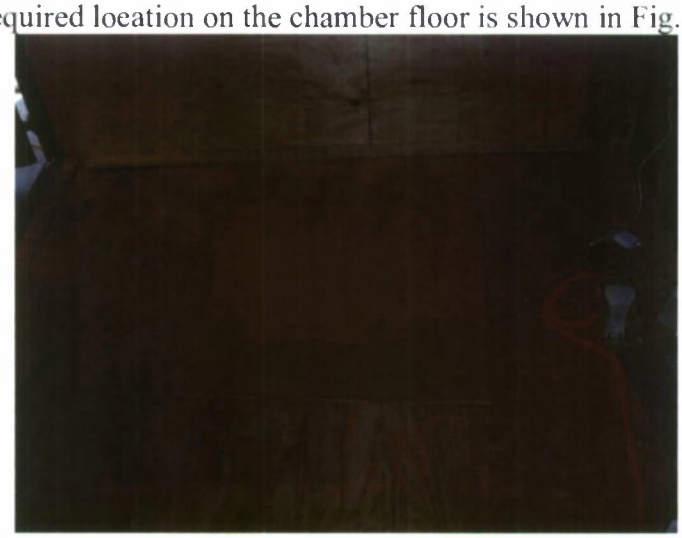


Fig. 13: Turntable Size and Location Relative to Front Wall

Since the motor must be located outside the chamber, two holes for the belt drive must be established through the screen room wall, and plywood-sheet metal and ferrite tile layers. Techniques to minimize the impact to isolation and tile absorption performance will be examined and the optimum method or combination employed.

The chamber will undergo characterization testing to compare radiated emissions performance against pre-compliance requirements, including isolation. Adjustments (isolation techniques mentioned above) will then be applied to improve overall performance.

Referenees:

1. D.R. Smith, W. Padilla, D.C. Vier, S.C. Nemat-Nasser, S. Schultz, "Negative Permeability from Split Ring Resonator Arrays," Lasers and Electro-Optics Europe, 2000, Conference Digest, 10-15 Sept 2000.

2. R.W. Ziolkowski, "Design, Fabrieation, and Testing of Double Negative Matematerials," IEEE Trans. Antennas Progat., Vol. 51, July 2003, pp. 1516-1529.

3. J.B. Pendry, A.J. Holden, D.J. Robbins, W.J. Stewart, "Magnetism from Conduetors and Enhaneed Nonlinear Phenomena," IEEE Trans. Mierowave Theory & Teehniques, Vol. 47, No. 11, Nov 1999, pp. 2075-2084.

4. J.B. Pendry and D.R. Smith, "Reversing Light with Negative Refraetion," Phys. Today, 57, June 2004, pp. 37-43.

5. ANSI C63.4-2003: Methods of Measurement of Radio-Noise Emissions from Low-Voltage Electrical and Electronie Equipment in the Range of 9 kHz to 40 GHz, American National Standards Institute.

6. CISPR 22: Information teehnology equipment, radio disturbance charaeteristics, limits and methods of measurement, Comité International Speeial des Perturbations Radioélectriques.

7. FCC, Part 15 (Radio Frequeney Devices): Subpart A: General, Subpart B: Unintentional Radiators, Subpart C: Intentional Radiators

8. M.I. Montrose, E.M. Nakauehi, *Testing for EMC Compliance, Approaches and Techniques*, Wiley & Sons, 2004

Red-tide Monitoring in San Luis Obispo Bay, CA Using LIDAR

Project Investigators:

Thomas J. Bensky and D. Rohan
Physics
California Polytechnic State University
San Luis Obispo, CA

ABSTRACT

Boundary layer structures caused by red tides and strong weather events in the ocean are measured using a LIDAR system, in parallel with an optical water-column profiler in San Luis Obispo Bay, California. Our data clearly reveals stark optical differences and layering structure in sea-water during red tide events, versus that during a clear tide. We also discuss data that reveals information about sediment motion in the water. Both this LIDAR project and the optical profiler were funded by C3RP, demonstrating a successful mating between separate C3RP efforts. We believe these efforts are consistent with ONR's goal of remote sensing at a safe distance.

1. BACKGROUND

California Polytechnic State University runs a marine research center, consisting of a 1 km pier at San Luis Obispo Bay, CA, which on the central California coast, approximately midway between Los Angeles and San Francisco. The pier extends over San Luis Obispo Bay, and is south-facing, so the usual mixing due to coastal storms is limited to mainly southern storms, which are less common for the California coast. This makes Avila Bay unique in that the water is relatively calm and changes occur over longer periods of time. We have installed a LIDAR facility at the end of this pier, and are able to optically access an approximately 30 ft deep water column of open ocean water.

2. LIDAR IN SAN LUIS OBISPO BAY

A 532 nm, 1 Watt YAG Laser, with a repetition rate of 20Hz is used as the LIDAR source. The laser is stationed at the pier, 1 km from the coast, directly over the open ocean. The laser is housed inside a control room and we are able to direct the beam out of this room, vertically into the water below. A single photon detector, prefixed by a narrow bandpass 532 nm filter, faces straight down, capturing photons scattered from the down welling laser beam. With the arrival of each photon, the detector provides timing pulses to a multi-stop, picosecond time analyzer (PTA), to capture the return times of photons relative to the initial laser shot to within an event-to-event resolution of 5 picoseconds. With this timing precision, subsurface features can be classified to within centimeters. Naturally, any concentrated particulate matter in the water will cause the laser light to scatter, some back to the photon detector. Of interest in this work is that several times a year, substantial layering occurs in the bay due to a "red tide" condition, when there is an increase in the number of algae near the surface. During this condition, the ocean water turns noticeably red or brown. Our current LIDAR research focuses on these events. Using an optical profiler in parallel with our LIDAR measurements, we also collect visible light samples down to a depth of ~25 feet. The profiler is an automated device that slowly lowers an upward-looking fiber optic cable into the ocean water. Light collected by the fiber is spectrally analyzed between 400 and 700 nm in increments of 1 nm, and logged at the collection depth. Data showing intensity collected at a particular wavelength, at a particular depth are shown below in Figures 1 and 2. Figure 1 was obtained during a red tide event on 11/15/2007 at 11:46am PST and 2b was taken during a clear tide on 11/17/2007 at 10:37am PST. The two data have sharply contrasting features. In particular, notice how abruptly the intensity of the signal drops across the entire visible spectrum during a red tide. In contrast, the clear tide graph shows a more moderate decrease with depth.

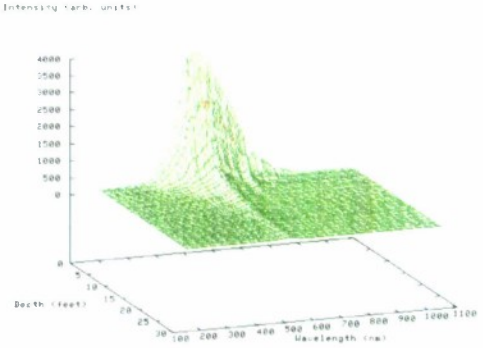


Figure 1: Spectra of light samples at depths from 0 to 25 ft during a red-tide event.

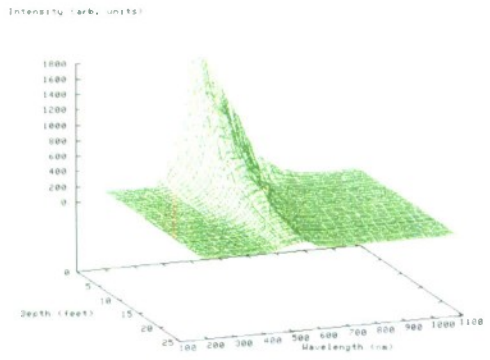


Figure 2: Spectra of light samples at depths from 0 to 25 ft during a clear-tide.

The optical profiler is useful in understanding how the LIDAR radiation, at 532 nm, interacts with the constantly changing optical characteristics of the sea water. By extracting intensity-versus-depth data from the plots in Figures 1 and 2 at 532 nm, for instance, we are able to obtain the instantaneous attenuation coefficient at the laser wavelength. These extractions are shown in Figure 3. As shown, the intensity of 532 nm light drops by 90% within the first 5 feet during a red-tide condition, requiring 25 feet to achieve this during a clear-tide.

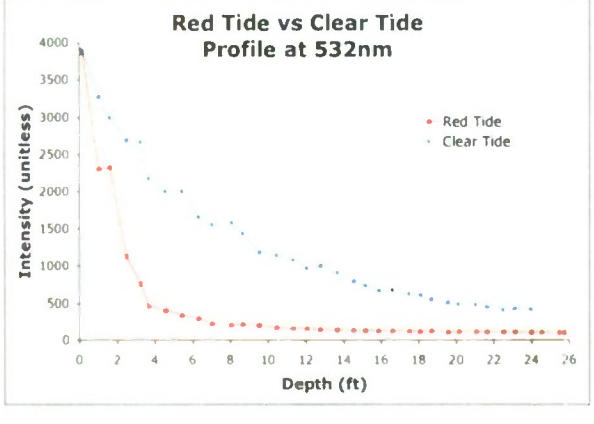


Figure 3: Intensity of 532 nm radiation as a function of depth, as extracted from Figures 1 and 2.

Attenuation of light in sea-water is known to be approximately exponential, so we apply a negative exponential curve fit, $y = Ae^{(-\beta x)} + B$. We find the β coefficient for the red tide data to be 0.5357/ft and 0.1617/ft for the clear tide. These are essentially the attenuation coefficients in both red and clear tide conditions, at 532 nm. The data observed under the same conditions by the LIDAR system (i.e. the arrival time of the returned photons) is shown below in Figure 4. In this Figure, the red curve was data taken during a red-tide event and the blue curve was for data taken during a clear tide.

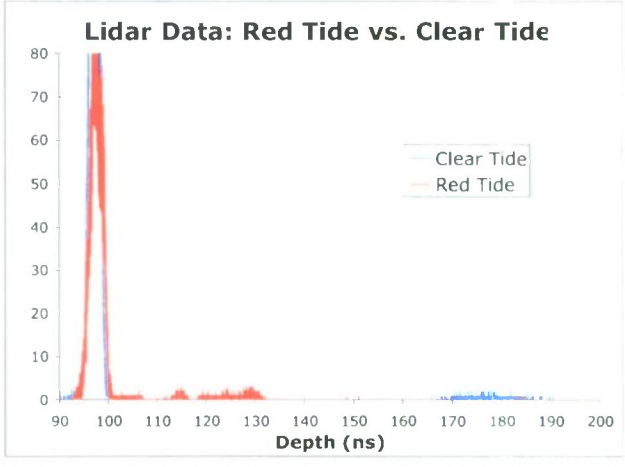


Figure 4: Return time of photons taken for laser shots over a one hour time span during red and clear tides.

Independent of water color, all of our LIDAR data shows a strong surface reflection at ~98ns. (There is no absolute significance to this time; it merely means that it takes 98 ns from the time the PTA is told to start its timing cycle for reflected surface photons to arrive back to our detector.) This surface reflection is not of immediate interest in this work, but sets the earliest time we expect to see photons arrive at our detector that interacted with the ocean water (it is a useful "time beacon"). Features plotted in red to right of this surface reflection (between 110 and 130 ns) were produced by the algae present during an intense red-tide event. In this case, the algae congregated close to the surface (within the first 10 to 15 feet) and presented the laser beam with a considerable scattering target. This feature is absent during a clear tide. The broad peak centered at 178 ns comes from the photons reflected from the ocean floor, and only appears during a clear tide (that is, it does not appear during a red-tide event). This indicates that the near-surface layer during a red-tide prohibits adequate light penetration to the sea-bottom.

An important observation to draw from Figure 4 is the fact that there are multiple features at different depths. This shows the power of LIDAR to characterize the water column below the laser. In a single data run (lasting about an hour) we obtain information about the surface, bottom, and the intervening water column. The system can measure multiple layers because of the

sensitivity of the detection and the high resolution of the timing system.

We further probe red tides with the LIDAR system by way of a retro reflector, which is a shiny metallic disc that can be manually lowered to any depths visited by the LIDAR beam, strongly reflecting laser light back toward the detector. Reflected photon intensity as a function of reflector depth during a red-tide event is shown in Figure 5.

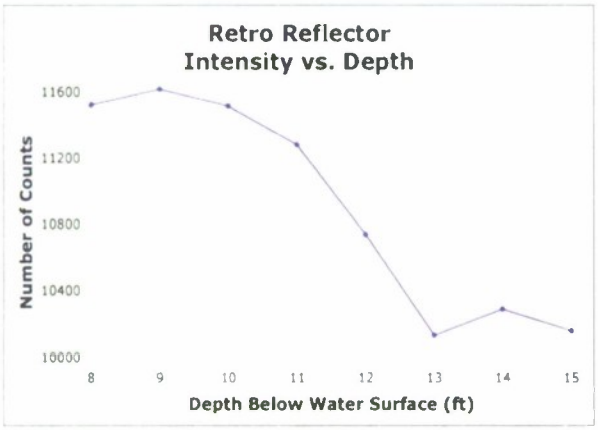


Figure 5: Laser pulse intensity reflected from a retro-reflector held at a given depth below the ocean surface.

This data shows a somewhat constant reflected laser intensity between 8 and 10 feet below the ocean surface. Between 10 and 13 feet, an abrupt drop in reflected laser intensity occurs, followed by an "elbow" at 13 ft, again returning a near constant intensity between 13 and 15 ft.

This data indicates relatively clear water to a depth of 10 feet, followed by a large amount of attenuation, presumably due to the algae bloom. The key feature is the sharp "elbow" at a depth of 13 ft, which we have observed for several different red tide events. We believe this coincides with the depth of the deepest algae boundary of the red-tide bloom. The laser intensity drops off quickly as a function of depth in the bloom (between 10 and 13 feet) until the retro reflector passes out of the layer and into clear water beneath the boundary. Thus we are able to use our LIDAR apparatus to estimate the height of a red-tide layer; in this case, an approximate 3-4 ft layer existed.

3. PARTICULATE LAYERING

The dynamic optical nature of the sea-water means that it is rare to see the same structures in the LIDAR return data over the short term (day by day). However, we are starting to see trends over longer periods of time. For instance, during calm winds fewer bottom-return photons are detected, compared times when the average wind speed increases to between 10 and 20 m/s. The wind appears to play a strong role in water mixing, leaving the water murky and "thick" in appearance when calm, and clearer and more "stirred" when strong.

As mentioned, our data always shows the strong surface reflection peak and another from bottom-reflected photons (if enough light can penetrate to the bottom and back). At other times though, we see multiple peaks throughout the water column, during clear water (not a red-tide condition) as shown in Figure 5. As verified visually with a video equipped remote operated underwater vehicle (ROV), the wind-driven currents appear to move localized sand layers in directions that are parallel to the sea-floor. These sand-layers provide ample scattering targets for the downwelling LIDAR pulse, sending a strong photon return signal back to the detector. Such a "sand storm" is shown in Figure 6.

Figure 6: Figure 6: A sand layer oberserved at a depth of approximately 5 m below the surface.

We note that with the ROV, the water was clear both above and below this sand-layer, indicating some localization of the sediment transport, which is consistent with the localization of the features seen in the LIDAR return data in Figure 5 between 120 and 140 ns. The feature between 180 and 220 ns is a bottom reflection. The large width indicates substantial sediment motion near the ocean bottom, where presumably the wave energy ultimately gets deposited as they approach the coast.

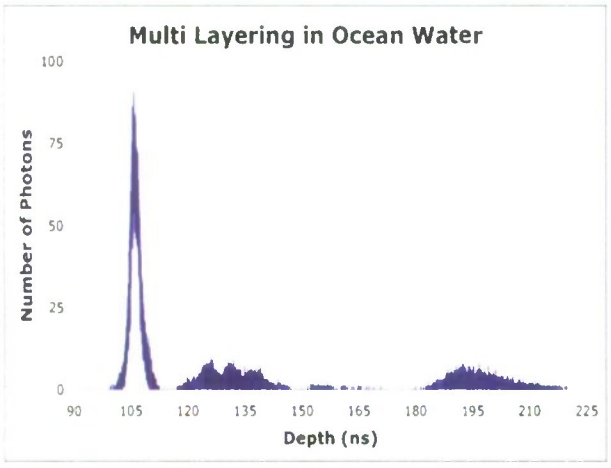


Figure 7: Layering due to particulates during windy (20 m/s) conditions.

4. CONCLUSIONS

Our LIDAR system coupled with an optical profiler has been used compare and contrast red versus clear sea-water conditions, from which we are able to extract red-tide layering depths and positions. Work continues that reveals current-driven sand events in a water column and sediment transport at the sea floor and at intermediate layers.

5. STUDENTS AND PRESENTATIONS

This work has resulted in senior projects and research work for the following undergraduate physics majors: Dov Rohan ('06), Chris Seck ('09) Derek Smith ('08), Collin Rodriquez ('08), Brenton Campbell ('06) and Brian Neff ('07).

The work was presented at the December 2007 meeting of the American Geophysical Union meeting in San Francisco, and at an American Physical Society meeting at UC Berkeley in November of 2007. A final presentation will be given on the International Laser Radar Conference in Boulder, CO, June 23-27, 2008.

Three-Dimensional Scaffolds for Mammary Epithelial Cell Growth

Project Investigators:

Robert Croekett, Biomedieal and General Engineering
Dan Peterson, Animal Science
Rafael Jimenez-Flores, Dairy Scienee
California Polyteehnie State University
San Luis Obispo, CA

Background: Original Project Objectives

The overall objective of this project was to develop techniques for creating complex, three-dimensional scaffolds supporting growth of functional mammary epithelial cells *in-vitro*. Mammary cell function is of interest from a fundamental cell biology perspective, but is also of potential commercial and strategic interest due to the potent secretory potential of this cell type. Secondly, the development of the tissue scaffold itself may be of commercial or strategic interest for tissue engineering applications.

The specific aims of the current funding cycle were three-fold:

- Define the optimal chemical (hormone) environment to promote maximal secretory potential in the MAC-T cell line that we are using in this bioreactor project.
- Compare the secretory activity (milk fat globules, proteins) and expression of lactation-specific gene products (mRNA, protein) of MAC-T cells grown on a conventional surface to those grown in the proprietary environment developed through this project.
- Expand the substrate development to include multiple micro- and macro-structural modifications and assess influences on cell function.

Summary of Project Outcomes

1. Additional funding from the California Agricultural Research Initiative (ARI) that has matched the funding acquired from the ONR (totaling ~$120,000) for continued work on the proteomic response of cells grown using the methodology developed herein.

2. Identification of previously unknown effects of hormonal treatments that influence mammary epithelial cell lipid distribution (manuscript in preparation for the peer-reviewed Journal of Dairy Science). **(Objective 1)**

3. Development and characterization of a potentially proprietary three-dimensional system for growth of secretory cells that may promote increased secretory activity as well as an increased ease of purification of secretory products. **(Objective 2)**

4. Development of potentially proprietary methodology for construction of novel extracellular matrices for support of tissue engineering applications, including but not limited to secretory cells. **(Objective 3)**

Results: Objective 1

The main limitation on secretory potential in the MAC-T cell line is the secretion of lipid droplets. There has been no description in the literature of secretion of lipids from a mammary epithelial cell line, and this may be a valuable and unique feature of this system, as well as of interest in terms of understanding basic secretory cell biology. As such, we investigated the synthesis of neutral lipid (NPL) in MAC-T cells which is only synthesized for the purpose of eventual secretion. MAC-T cells were grown on flat plastic surfaces as a baseline for comparison of hormonal influences alone. In a 3 x 3 factorial arrangement, the effect of growth

hormone (GH; 0, 10, 100 ng/ml), and prolaetin (PRL; 0, 5, 10 µg/ml) on the lipid composition of MAC-T was examined using thin layer ehromatography. GH had a signifieant effeet on the proportion of NPL, cholesterol, phosphatidyleholine (PC), phosphatidylinostitol (PI), and sphingomyelin (SM) (table 1). Only the proportion of SM was affected by PRL eoneentration. The interaetion between GH and PRL treatments was not signifieant. The relationships between the different lipid elasses and hormone treatments appeared to be eomplex. In partieular, the proportion of NPL was highest in response to the 10 ng/ml GH, while 100 ng/ml GH led to intermediate levels.

GH (ng/ml)	0			10			100				
PRL (µg/ml)	0	5	10	0	5	10	0	5	10	Effect	SEM
NPL	11.07	10.13	11.22	17.19	18.57	18.66	15.00	12.13	13.04	GH	0.76
Cholesterol	21.15	18.79	19.77	25.28	24.45	24.11	24.80	23.63	24.21	GH	0.69
PE	6.65	6.97	6.57	7.05	6.74	6.43	5.84	6.75	7.18		0.20
PC + PI	6.98	7.28	7.10	6.34	5.97	5.98	5.24	5.52	6.34	GH	0.20
Unknown 2	10.96	9.61	10.76	7.14	7.19	6.36	6.53	7.79	6.95		0.55
SM (0.141)	24.45	24.29	24.10	25.16	24.60	25.32	21.40	22.59	22.02		0.77
SM (0.077)	6.61	13.35	8.58	3.67	3.69	3.90	9.46	11.80	11.45	GH,PRL	0.26
SM (Combined)	31.07	37.64	32.68	28.83	28.29	29.22	30.86	34.38	33.47	GH	0.64
Unknown 1	5.17	4.27	5.07	2.60	2.81	2.90	5.44	4.98	4.41	GH	0.30
Origin	6.95	5.32	6.84	5.58	5.97	6.34	6.30	4.82	4.41		0.80

Table 1. The mean pereentages of all lipid fraetions aeross all treatments are presented. The standard error of the means (SEM) for eaeh lipid fraetion aeross all treatments was also ealeulated. The GLM was used to analyze the effeet of GH, PRL and the interaetion between GH and PRL on the pereentage of eaeh lipid fraetion within eaeh sample. The dependent variables that had a signifieant effeet on the pereentage of eaeh lipid fraetion are shown. GH eoneentration had a signifieant effeet ($p \leq 0.05$) on the proportion of NPL, eholesterol, PC+PI, SM (0.077), SM (eombined). PRL, had a signifieant effeet on SM (0.077). The interaetion between the GH and PRL did not have a signifieant effeet on any of the lipid fraetions.

Continued work is aimed at defining the proteomie response (effeets on global protein expression) of mammary epithelial eells to the different hormone treatments previously deseribed. To date, we have identified proteins expressed in MAC-T cells as well as those present in the media, and are eurrently developing methodology to identify alterations due to treatment (hormonal or physical in nature) by two-dimensional gel electrophoresis **(supported by ARI funding).**

Results: Objective 2

Our initial attempts at growing eells in the environments developed as part of this projeet have met with mixed sueeess. The exaet nature of the environment will not be deseribed in detail due to their potential proprietary nature. Initial results have shown the environment to be biologieally eompatible, leading to eell habitation and proliferation (see Figure 2), though not to the extent that was expected.

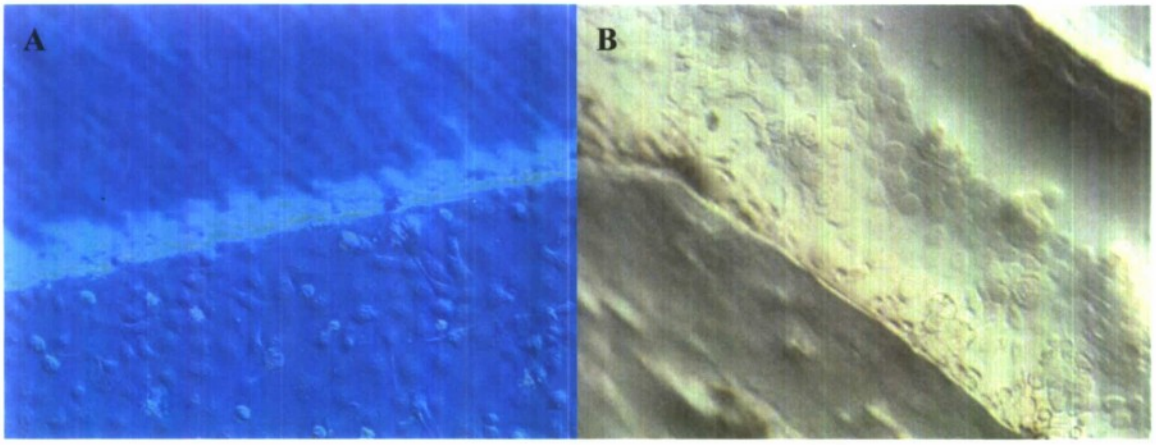


Figure 1. MAC-T cells grown on extracellular matrix developed in this project. A: eells grown on material in simple geometry as a proof of eoneept. B: eells grown on material in initial complex gcometry targeted for further development.

Based on our trials, the limitation of the biomaterial developed in this projcct is not relatcd to it's chemical structure, as cell proliferation is not limited in comparison to growth on flat plastic (Figure 1A). However, when grown in eomplex three-dimensional cnvironmcnts (Figurc 1B), eells were of limitcd proliferativc capaeity (data not shown). Possible limitations of thc substrate were postulated to include permeability characteristics of the material itself. As such, we engaged in analysis of the permeability of the material to complex biological molecules such as protcins. This investigation has led us to systematically investigatc thc permeability of the material with respeet to eomplex protein mixtures.

Results have indieated that different formulations of thc biomatcrial have drastieally different permcabilities for different proteins (Figure 2). Interestingly, these differenees do not seem to be rclatcd only to protein size, as some larger proteins diffuscd readily, while somc smallcr proteins diffused to a mueh lesser extent over a 24h period. The nature of and basis for these differenees arc currently being explored as part of the ARI-funded project. Further understanding of these pcrmcability characteristics is expected to yield a greater understanding of thc biocompatibility of the matcrial developed.

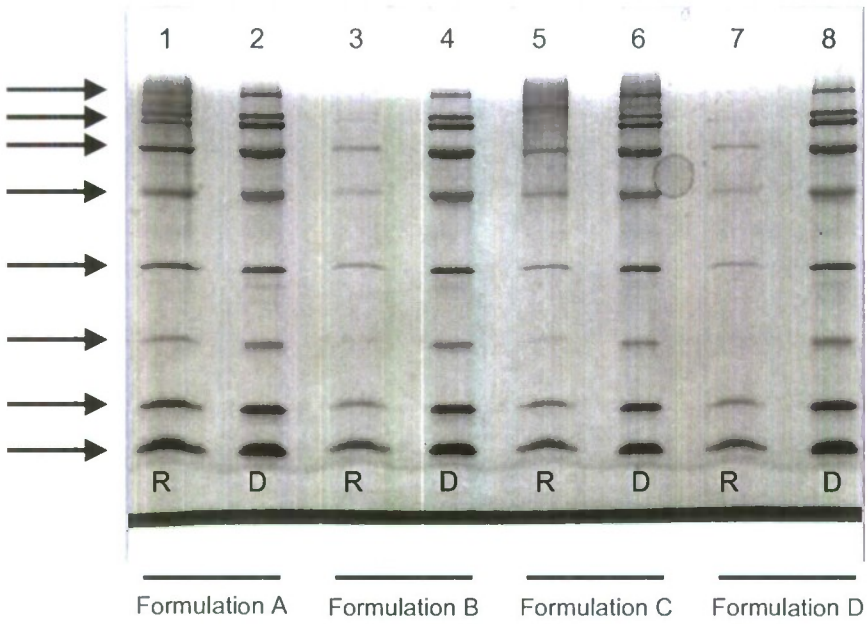


Figure 2. Permeability trial for different formulations of extracellular matrix biomaterial. Protein mixture was separated from water by the biomaterial formulations A-D and allowed to diffuse for 24 hours and then separated by size (smaller at the bottom, larger at the top). D denotes the donor chamber, while R denotes the receiving chamber. Arrows indicate proteins that diffused to different extents dependent on matrix formulation.

Results: Objective 3

Detailed descriptions of our progress on Objective 3 will be excluded from this report, as they are proprietary in nature. Briefly, additional three-dimensional structures are being explored, as well as novel means of introducing cells to these structures, and are being analyzed for their effects on cell survival, cell morphology, and cell-cell interactions. The effects of multiple cell types in a given environment are also being explored as we continue to investigate the biocompatibility of the developed material.

Conclusions

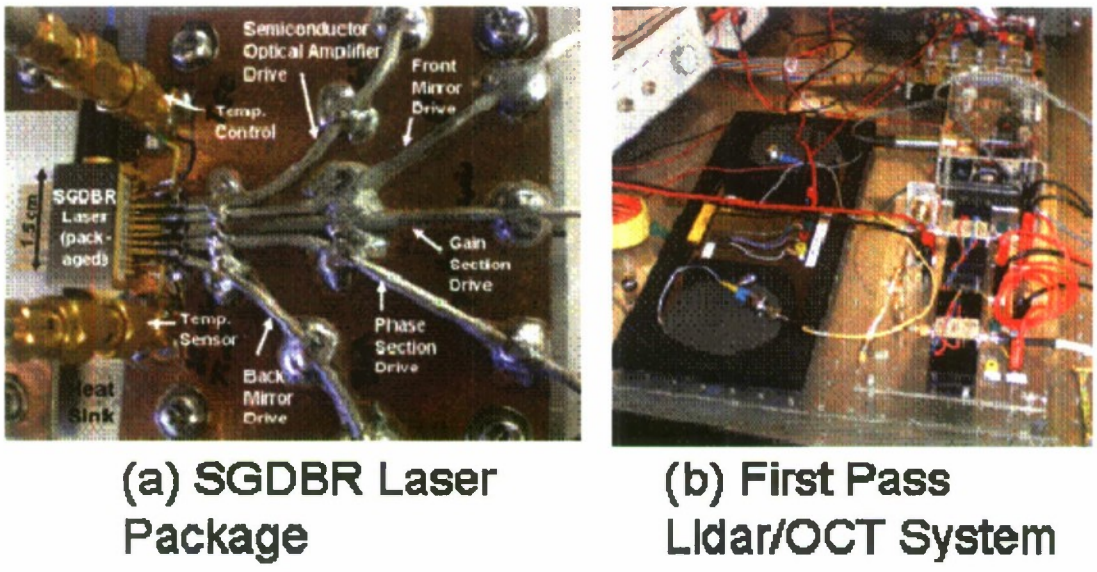
Overall, this project has led to new information pertaining to hormonal influences on secretory cell function, and this work is continuing under funding from the CSU Agricultural Research Initiative. This project has also led to the development of a novel cell culture environment that has proved to be biocompatible, and analysis of the molecular properties of the material is in progress. Finally, novel complex environments for cell growth continue to be developed in the interest of defining physical influences on tissue growth and function in artificial environments.

SG-DBR Widely Tunable Lasers for LIDAR and Optical Coherence Tomography

Project Investigator:

Dennis Derickson
Electrical Engineering
California Polytechnic State University
San Luis Obispo, CA

Executive Summary:



(a) SGDBR Laser Package

(b) First Pass Lidar/OCT System

Figure i: (a) This picture shows the packaged Sampled Grating Distributed Bragg Reflector Laser that was designed for LIDAR and OCT applications. Part (b) shows the optical and electronic hardware used to interface to the laser for LIDAR signal processing.

1. For the first phase of this project we assembled a LIDAR/OCT system utilizing the SGDBR laser as pictured in figure A. Initial LIDAR and OCT results were obtained. Initial results were presented at the BIOS 2008 conference in San Jose, CA.

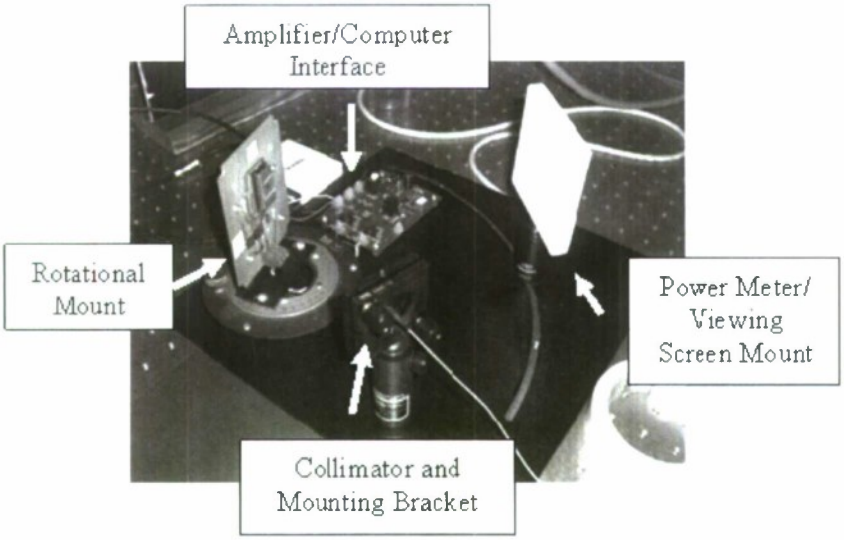


Figure B: This photograph shows an experimental set up used to measure the characteristics of a MEMS based optical beam steering mirror.

2. We investigated a MEMs steering mirror for steering the optical beam in LIDAR applications. Figure B illustrates the assembly that was built and tested. The technology looks promising but several device anomalies like hysteresis make the electronic drive control much more complicated for LIDAR applications.

This work was sponsored by the Department of the Navy, Office of Naval Research, under Award # N00014-06-1-1111.

Major Sections in the Report

1. **Overview of Project Definition:** The first section of the report quickly reviews what was being proposed in this project in January of 2007.
2. **SGDBR Lasers and Their Wavelength Tuning:** The second section of the report goes through a detailed study of the SGDBR laser and how it can be used to construct the linear wavelength ramp needed for chirped radar and Optical Coherence Tomography (OCT) applications.

3. **MEMs Beam Steering:** The final section of the report covers an evaluation of a MEMS steering mirror that could be used for beam steering.

1. Overview of the Project Definition

LIDAR (Light Detection and Ranging) system architectures and applications that are enabled by single-chip wavelength tunable lasers will be developed in the proposed work. These new single chip lasers are unique in that wide-wavelength tunability (60 nm at 1550 nm center wavelength), high-speed frequency switching times (1ns), and high-speed amplitude modulation (5 GHz) functions are all integrated on the same semiconductor chip. The goal of this work is to adapt these lasers for applications in LIDAR and OCT. The first stage of proposed research (the subject of this proposal) is to find architectural solutions that utilize the strengths of this laser class and mitigate their weaknesses (table 2 below).

2. SGDBR Lasers and Their Wavelength Tuning:

The application of Sample Grating Distributed Bragg Reflector (SG-DBR) lasers in Optical Coherence Tomography (OCT) and LIDAR applications is explored. Specifically, this report involves using SG-DBR lasers to improve upon the source component of current LIDAR and OCT architectures. The report investigates the Chirped Laser Architecture and making changes to the architecture to accommodate an SG-DBR laser.

This report involved investigating SG-DBR lasers. More specifically, it investigated the tuning characteristics of the Agility 1107 CW SG-DBR Widely Tunable Laser (Agility 1107) and its applicability as a LIDAR/OCT system source. After investing SG-DBR lasers, the report concentrates on developing an open-loop tuning algorithm using the Agility 1107 SGDBR laser in a custom developed package. Aspects such as repeatability and hysteresis are discussed in strategies for open-

loop tuning. Developing the tuning algorithm involved overcoming particular device hurdles. Specifically, the report addressed the wavelength repeatability of the device when overcoming hysteresis.

Finally, the report tested the architecture employing SG-DBR lasers and proved that the laser can be used in LIDAR/OCT applications. Using a fiber as a test device, the systems function was verified by obtaining a linear change in phase and a constant magnitude.

CHAPTER 1 Introduction

1.1 Overall Objective

This report presents an open-loop tuning method for Sampled-Grating Distributed Bragg Reflector (SG-DBR) Widely Tunable Lasers to be used in Optical Coherence Tomography (OCT) and LIDAR. To orient the reader on this overall objective it will be dissected into its individual parts.

The first section of this report will orient the reader on OCT techniques and applications. OCT is a noninvasive tomographic imaging method. Due to the nature of the wavelengths involved in OCT, the imaging method has been proven useful in imaging micro-particles like organic tissue and glucose [1]. This report will address improvements which can be made on current OCT technology by improving upon the source component of current OCT architectures.

The second section of this report will describe a new type of widely tunable single chip laser called the SG-DBR laser. This laser is new to the market and was originally intended for telecommunications applications. After a detailed discussion on the workings of the SG-DBR laser, the report will outline why this laser could be so valuable in OCT applications. Like any new potential solution this will offer advantages and disadvantages compared to other OCT sources. The discussion presented in this report will address both the pros and cons of the solution.

This report will concentrate on several features of the laser source in an OCT architecture, particularly, its fast wavelength switching capabilities. This report

presents an open-loop tuning algorithm which allows a system to take advantage of the fast wavelength switching characteristics of the SG-DBR laser.

In the process of developing a tuning algorithm several problems were encountered due to the nature of the laser. This report presents a solution to these problems in the process of developing an open-loop tuning algorithm.

Finally, the OCT measurement technique will be demonstrated using the SG-DBR laser as the source component. Overall, this report should show the feasibility of using SG-DBR lasers as the source component of an OCT system and improve on current OCT technology.

CHAPTER 2 Optical Coherence Tomography

2.1 What is Optical Coherence Tomography?

Optical Coherence Tomography (OCT) is an imaging technique that produces cross sectional images based off information from optical reflectivity. It is based on the principle of low-coherence interferometry where distance information concerning various micro-structures is extracted from time delays of reflected signals. An OCT system functions similarly to a time-domain reflectometer [2].

The OCT imaging technique can be used to obtain high-resolution cross-sectional images within both transparent and semi-transparent environments such as organic tissue and thin films [3]. OCT has various uses beyond cross-sectional imaging such as glucose monitoring [1]. The main benefit of OCT is that it is noninvasive. An informative high-resolution image of organic tissue can be obtained and analyzed without surgical procedure, making the technique very appealing in medical practices.

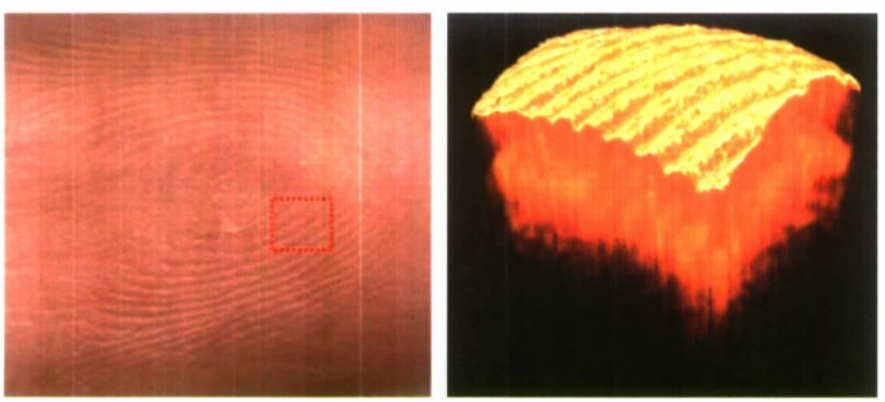


Figure 2.1 - An example of an OCT image taken by Thor Labs. (a) The picture to the left shows a human finger print. (b) The picture to the right shows an OCT image of the human finger print.

Figure 2.1 (a) shows an example device under test (human fingerprint). Figure 2.1 (b) shows an example of the high resolution image which can be obtained using OCT. Notice the level of detail of the finger print in Figure 2.1 (b). Since OCT can be performed using wavelengths within the near IR Band (1200nm – 1600nm), it is capable of penetrating millimeters into the skin and capable of resolving objects on the micrometer scale [2,14].

OCT is a form of Optical Vector Network Analysis (OVNA). OVNA refers to measurement of both the magnitude and phase response of a device under test. Unlike a scalar analysis which measures only magnitude response, vector analysis includes signal delay information with magnitude and phase allowing for both magnitude and range measurements. Figure 2.1 illustrates the function of an OVNA system.

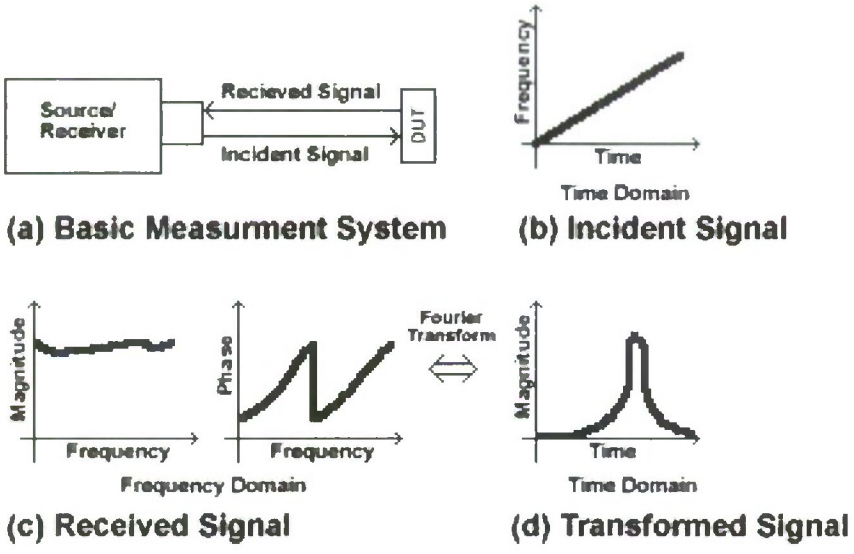


Figure 2.2 - The OVNA measurement system. (a) The basic OVNA measurement system consists of a source/receiver and device under test (DUT). (b) Incident signal used from source is a linear ramp of frequency versus time. (c) The received signal contains both magnitude and phase data. (d) Transformed magnitude and phase data yields a time domain response.

In Figure 2.2 (a) the source radiates a device under test (DUT) with a specified signal. In the case in Figure 2.2 (a) a linear ramp of frequency versus time is used and is shown in Figure 2.2 (b). The receiver then collects the signal reflected from the DUT. The receiver collects both the magnitude and phase data of the reflected signal, this is shown in Figure 2.2 (c). The magnitude and phase data in frequency domain can be converted into the time domain via the Inverse Fourier Transform. Performing the transform yields Figure 2.2 (d). The example in Figure 2.2 shows how measurements in the frequency domain can yield information in the time domain.

2.2 Why OCT?

OCT frequencies range from that of visible light all the way down to 200 THz. It's been shown that these frequencies are useful in the imaging of organic tissue such as skin [17]. These frequencies do not ionize atoms as do x-rays, and therefore leave tissue unharmed. The noninvasive nature of the measurement scheme makes it a fast low cost alternative to x-rays or CAT scans in specific medical applications. However, the limitations to OCT in medical applications when compared to x-rays or CAT scans is the penetration depth [2].

2.3 OCT Architectures

There are two main measurement architectures used in OCT applications: White Light Interferometry and the Chirped Laser Technique.

2.3.1 White Light Interferometry

White Light Interferometry has been a well established technique for measuring the magnitude and delay response of a device under test. The technique has been refined and become useful in imaging different types of tissue [14]. The architecture

for a White Light Interferometry System can be seen in Figure 2.3.

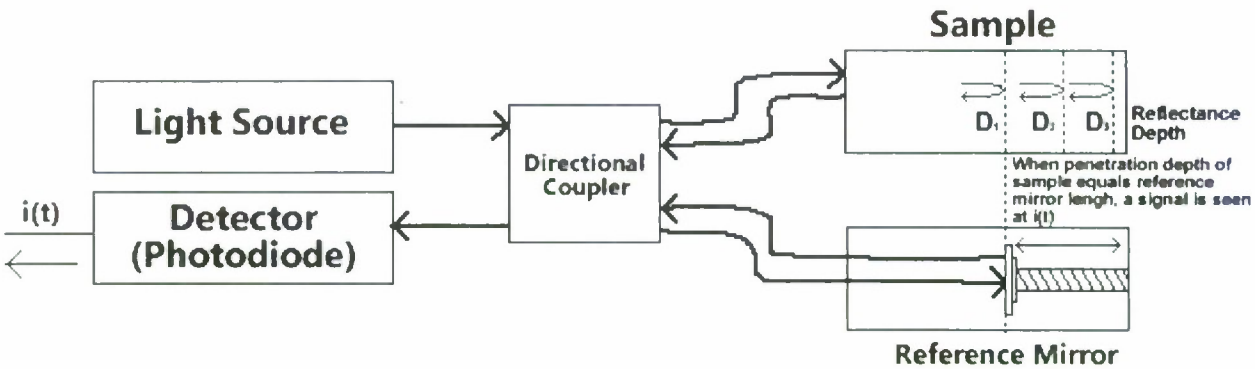


Figure 2.3 - The architecture of a White Light Interferometry System. The bold connection arrows signify optical paths. A current i(t) is seen when the path from the sample reflection is equal to the path from the mirror.

Figure 2.3 shows light emitted by the "white" light source being fed into a 50/50 directional coupler. The spectrum of a white light source is shown in Figure 2.4. An equal amount of light is sent to the sample and a reference mirror. The configuration between the directional coupler, sample, and reference mirror is known as a Michelson configuration which is related to the Michelson Interferometer [15].

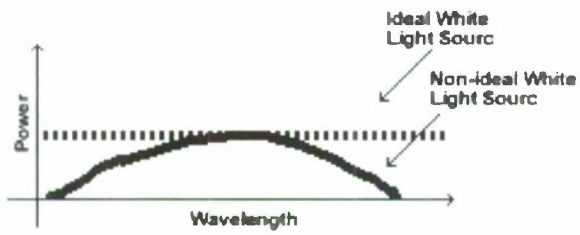


Figure 2.4 - The spectrum of a white light source.

In Figure 2.3 the mirror acts as a reference point for the light reflected from the sample. When the penetration distance into the sample is unequal to the reference distance of the reference mirror, a DC signal to appear at i(t) of Figure 2.3. The

signal appearing at i(t) in Figure 2.3 is shown in Figure 2.5. When the distances are identical, interference between the two signals occurs and an AC signal appears at the output of the detector. The signal at i(t) in Figure 2.3 is related to the source bandwidth through time resolution of interference patterns (see Figure 2.5). For example, the time resolution of the interference patterns in Figure 2.5 for a 10THz source would be one picosecond. The interference patterns generated in Figure 2.5 represent the reflectivity of the sample. Hence, magnitude and phase are both processed on the fly in the time domain.

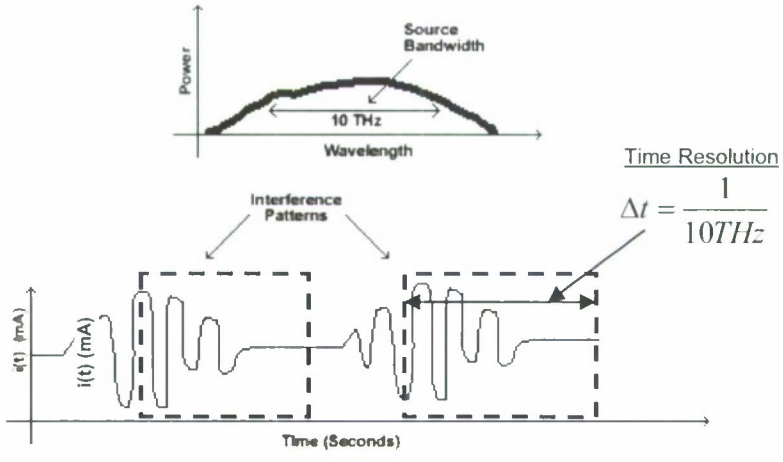


Figure 2.5 - Interference patterns from the White Light System. The interference patters generated using white light architecture are signified by the dotted boxes.

The benefit of using the technique shown in Figure 2.3 is that the entire measurement is done in the time domain, so no post-processing is necessary. The disadvantage of this technique is that mechanical parts are necessary to move the mirror. Therefore, high-resolution real-time images requiring a fast measurement and data acquisition scheme are harder to obtain [4].

The system shown in Figure 2.3 has signal-to-noise ratio which is inversely proportional to the detection bandwidth of the system [3]. Fast scanning raises the detection bandwidth of the system[3]. The goal of an OCT system is to have both a high scanning speed and a high signal-to-noise ratio.

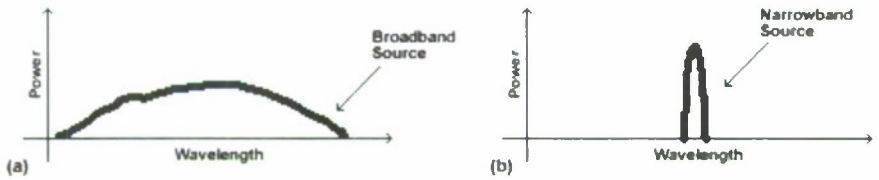


Figure 2.6 - Broadband/Narrowband Source Comparison. (a) Illustration of the power spectrum of a broadband source. (b) Illustration of the power spectrum of a narrow band source.

Figure 2.6 illustrates the power spectrum of both a broadband source and a narrowband source. The source seen in Figure 2.6 (a) is the source used in the White Light Architecture. The time resolution of the source is proportional to the inverse spectral width of the source (shown in Figure 2.5 (a)). Therefore, the time resolution for a system using the source seen in Figure 2.6 (a) is smaller than for the source seen in Figure 2.6 (b).

Since the White Light Architecture requires one "white" light source seen in Figure 2.6 (a), it would take a great amount of power to have the source yield a high signal-to-noise ratio at all frequencies. More specifically, the source seen in Figure 2.6 (b) has all its energy concentrated over a smaller area than the source seen in Figure 2.6 (a). Therefore, the source seen in Figure 2.6 (b) would yield a higher signal-to-noise ration than the source in Figure 2.6 (a). The signal-to-noise ratio comes into play with both more accurate images and a larger penetration depth and a larger signal to noise ratio [3]. This report will explore another way to obtain the same time

resolution as a broadband source by using a narrowband source.

2.3.2 Chirped Laser Interferometry Architecture

The Chirped Laser Interferometry Architecture can be seen below in Figure 2.6.

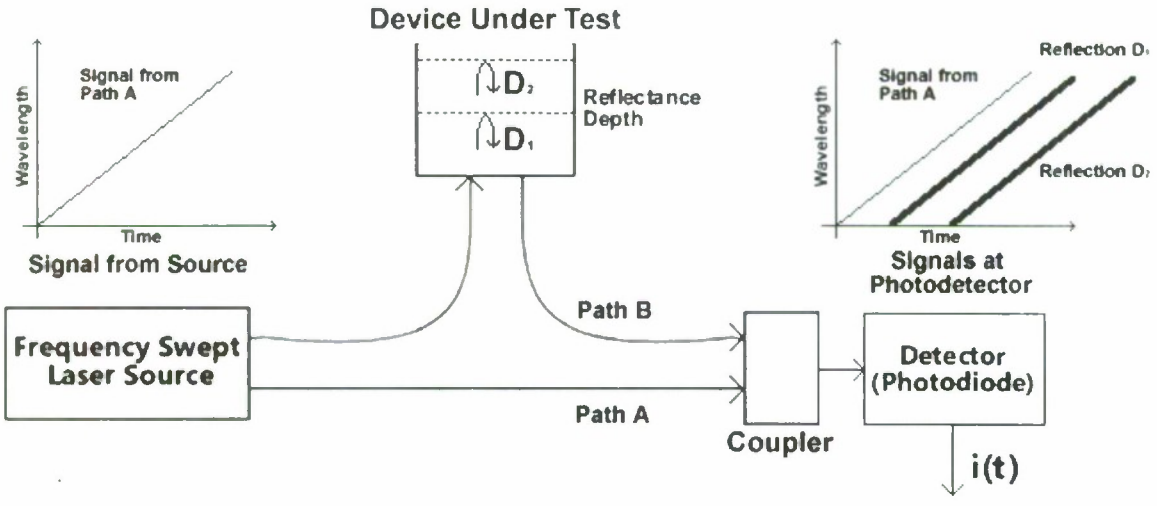


Figure 2.7 - Chirped Laser OCT/LIDAR Architecture.

The main difference between the Chirped Laser Architecture of Figure 2.7 and the White Light Architecture of Figure 2.3 is within the source component of the system. The White Light Architecture of Figure 2.3 consists of a constant "white" light source seen in Figure 2.6 (a), while the Chirped Laser Architecture of Figure 2.7 employs the narrowband source component seen in Figure 2.6 (b) ramped in the frequency domain. The system collects imaging information by linearly ramping the narrowband source of Figure 2.6 (b) in time with respect to wavelength, and comparing the difference between incident wavelength magnitudes emitted from the source to the reflected wavelength magnitude from the device under test. Figure 2.7 shows an illustration of this frequency comparison by showing the reflected signal from the DUT compared at the photodetector.

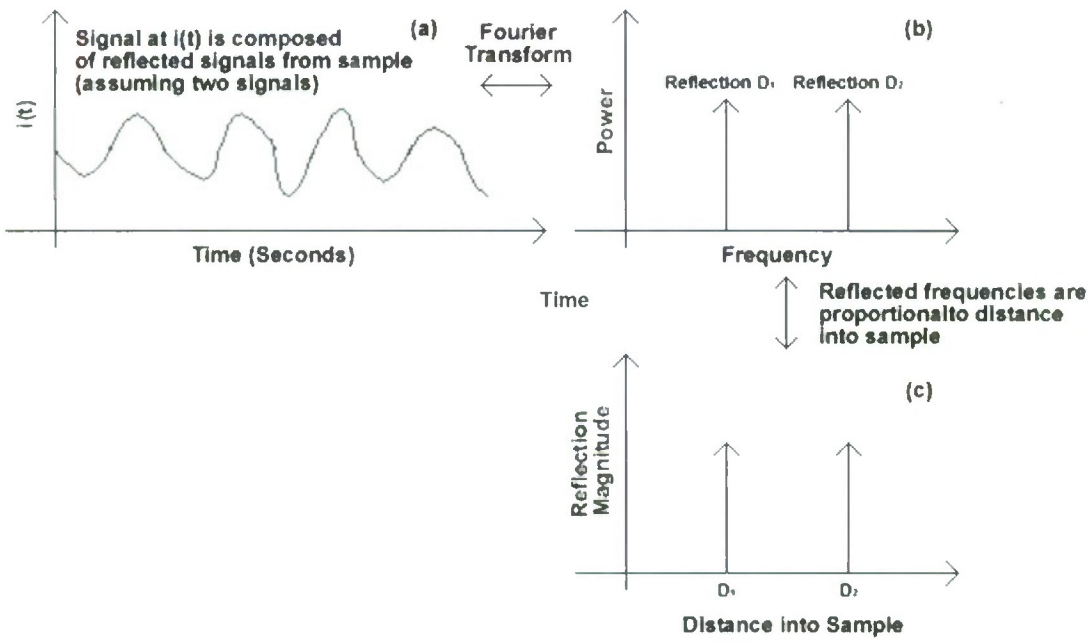


Figure 2.8 - Signal analysis of Chirped Laser Architecture. (a) An example of a signal i(t) emitted by the photodetector in Figure 2.7. The signal is assumed to be composed to signals from two reflections. (b) An example Fourier transform of the signal emitted from the photodetector in Figure 2.7. (c) The distance within the sample is proportional to the frequencies of the signals emitted by the photodetector.

In order for this comparison to be accurate, the frequency spacing must be constant meaning the ramp must be linear. The reason for this is that in the Fourier Transform (2.1), k is an integer, meaning that the spacing between frequencies is f_0. Therefore, in order for the chirped laser technique to work accurately, evenly spaced frequency steps must be used [16].

$$X[k] = \sum_{k=0}^{\infty} x[k] e^{-j2\pi k f_0 t} \tag{2.1}$$

There are a couple advantages to the Chirped Laser Architecture over the White Light architecture. The first advantage is that since the laser's power only is

distributed over a narrow frequency band, the signal-to-noise ratio of the interfering signal at the detector is stronger than those in the white light case. Second, since the white light architecture needs to sweep a mirror, it takes longer to collect data and involves the motion of mechanical parts. However, SG-DBR widely tunable lasers can switch wavelengths in the nanosecond range [6]. Therefore one can gather more data in a smaller amount of time using the Chirped Laser approach and this leads to real-time OCT applications. A similar approach to the Chirped Laser Technique will be used in this report in order to explore the capabilities of SG-DBR widely tunable laser in OCT.

2.4 Modulation Phase Shift Method for Group Delay Measurements

The third method utilized in this work to measure group delay is shown in Figure 2.9. Figure 2.9 is an illustration of the Stepped Ramped Wavelength Architecture. Figure 2.9 shows the output of a tunable laser source modulated and fed into a DUT. The tunable laser is being swept in a tuning ramp fashion. The reflected signal (Signal (b) in Figure 2.9) from the sample is fed into a photodiode where the optical signal is converted into an electrical signal (Signal (c) in Figure 2.9). The current emitted from the photodiode has a phase difference compared to the initial modulation source since the signal has been affected by the DUT. The system basically uses Signal (a) of Figure 2.9 and Signal (c) in Figure 2.9 to obtain phase and magnitude measurements.

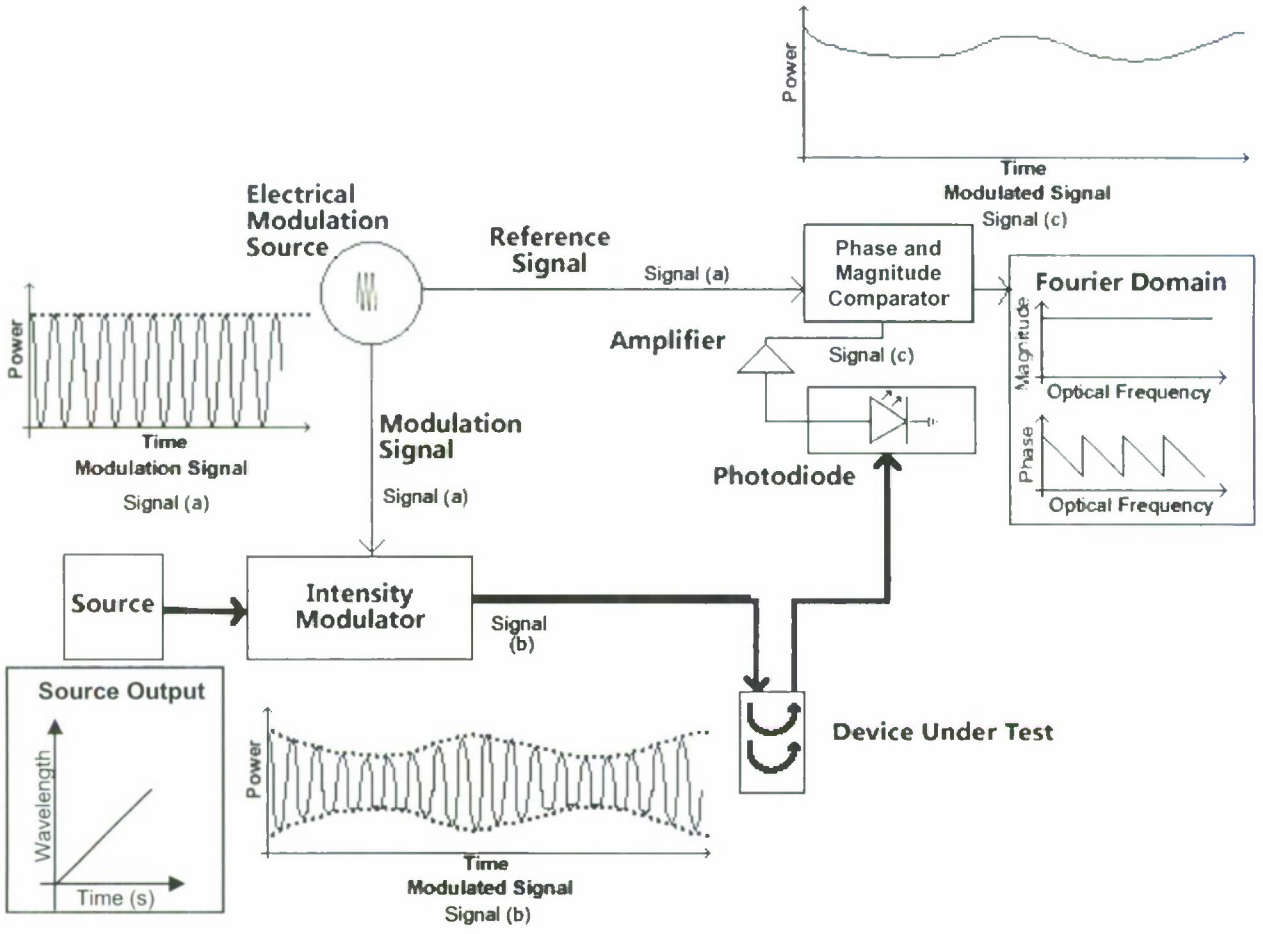


Figure 2.9 - Illustration of the Stepped Ramped Wavelength Architecture. The bold connections signify optical paths. Signal (a) represent the modulation signal. Signal (b) represents the source signal modulated and sent to the device under test. Signal (c) represents the signal from the photodetector containing information about the DUT.

The phase and magnitude of the current from the photodiode in Figure 2.9 is then compared to the phase of the modulating source. Since the system in Figure 2.9 allows for the collection of both magnitude and phase measurements, it is useful for collecting data in the Fourier Domain and making group delay measurements. Group delay is the measurement of change in phase with respect to change in frequency and is defined in (2.2).

$$Group\ Delay = \frac{\Delta Phase(radians)}{\Delta Frequency(radians/\sec ond)} \tag{2.2}$$

It must be noted that the system in Figure 2.9 measures change in group delay rather than delay itself. This measurement of change in group delay is due to the fact that the optical frequency is being ramped in Figure 2.9, not the modulation frequency. The expression for calculating change in group delay with respect to changing optical frequency (not the modulation frequency) is shown in (2.3).

$$\Delta\tau_{\lambda} = \frac{\Delta Phase}{\Delta Optical\ Frequency} \tag{2.3}$$

The slope of the phase in Figure 2.10 is constructed by finding the change in group delay at all the wavelengths within the tunable source's range. This method of measuring change in group delay will be used to verify proper functionality of the OCT system. The system in Figure 2.10 shows Stepped Ramped Wavelength Architecture using a fiber optic cable as the device under test.

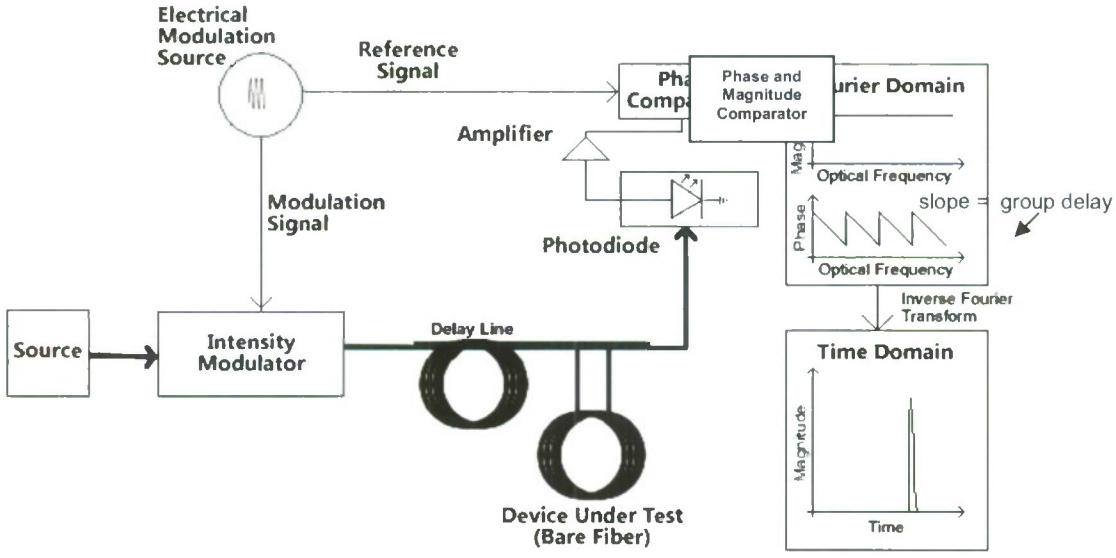


Figure 2.10 - Illustration of Stepped Ramped Wavelength Architecture. The bold lines represent optical paths.

Notice, in the Time Domain processed data of Figure 2.8 there is a spike. This spike represents the group delay within the fiber. Knowing the speed the signal is traveling through the fiber, one can use this delay to calculate the length of the fiber.

2.5 Current System

The OCT system used in this report is a variation of the Stepped Ramped Wavelength Architecture shown in Figure 2.7. Figure 2.11 shows what will also be referred to as the Stepped Ramped Wavelength Architecture. The main difference between Figure 2.7 and 2.11 is that a network analyzer is being used to collect magnitude and phase data in Figure 2.11. However, the system in Figure 2.7 is fundamentally the same as the system seen in Figure 2.11. The goal of this report is to improve upon the current OCT technology by employing a widely tunable laser source which has very fast wavelength switching capabilities (nanoseconds) [6]. The design being explored in this report is designed to leverage the laser source's (SG-DBR Laser) strengths against its weaknesses.

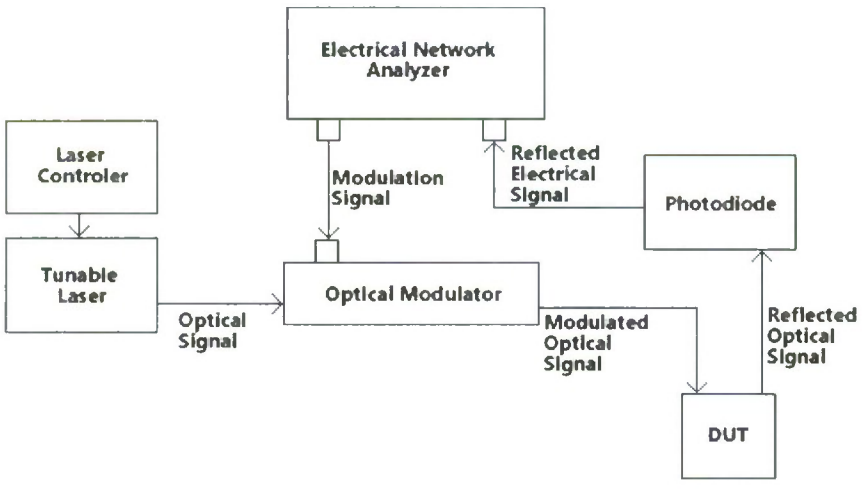


Figure 2.11 - Illustration of the Ramped Wavelength Architecture.

There are a few additions to the diagram in Figure 2.11 which were not shown in Figure 2.7. Mainly, the electrical network analyzer will be used to gather magnitude and phase data from the reflected signal. It will also be used to generate the modulation signal used to modulate the tuning ramp. The other changes to the architecture will be discussed in Chapter 6.

2.6 Development

In designing an OCT system the chirped laser technique has been shown to be more promising due to faster switching capabilities and having no reliance on mechanical parts. This report will explore changes made to the source component of the Chirped Laser Architecture that can potentially improve the overall OCT System.

It is important that the frequency steps are of equal spacing when constructing a tuning ramp due to the fact that a ramped laser architecture depends on the Inverse Fourier Transform to processes the frequency domain measurements into a time domain response.

The next chapter discusses a unique laser known as a SG-DBR Single-Chip Widely Tunable Laser (SG-DBR). The Stepped Ramped Wavelength Architecture will take advantage of the laser's switching speed in order to lead to the development of OCT system. The scope of this report is to set up a working prototype for the development of an OCT system which employs. SG-DBR lasers as the source component.

CHAPTER 3: Widely Tunable SGDBR Lasers

3.1 Introduction to Diode Lasers

This chapter will provide the reader with background information about SG-DBR lasers and the details about the specific SG-DBR laser used in this report. That background will be followed by an in depth view of the laser's tuning mechanism and a glance at some of the underlying physics concepts associated with the SG-DBR laser. The goal of this chapter is to familiarize the reader enough with the laser to understand the application performed in the following chapters.

3.1.1 Choice of Laser

When choosing widely tunable diode lasers, there are two main types: single-chip (otherwise known as Monolithically Integrated (MI)) and External Cavity Lasers (ECL). The laser used in this report is an MI SG-DBR widely tunable laser. There are two major advances of MI tunable lasers over ECLs:

1. Size – MI widely tunable lasers are much smaller than external cavity lasers.
2. Speed – MI widely tunable lasers have a much faster wavelength switching speed than do external cavity lasers.

3.2 Agility 1107 CW Widely Tunable Laser

The laser used in this research is the Agility 1107 Continuous Wavelength Single-Chip SG-DBR Widely Tunable Laser (Agility 1107). The Agility 1107 can be seen in Figure 3.1.

Figure 3.1 - The Agility 1107 laser. (Left) Packaged version of the Agility 1107 Widely Tunable Laser [13]. (Right) The image of the package version the Agility 1107 while used in the laboratory.

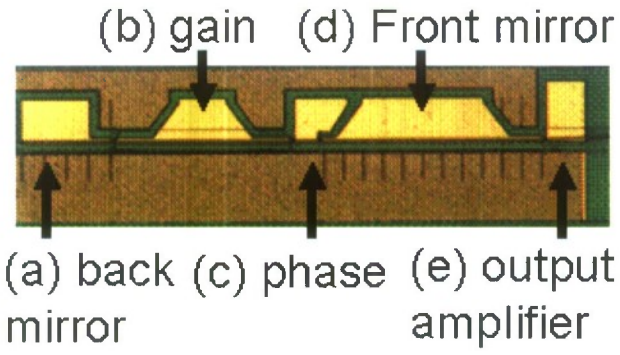


Figure 3.2 - Illustration of SG-DBR laser elements. An illustration of a close-up overhead view of an SG-DBR Widely Tunable Laser. Sections (a), (c), and (d) are involved in wavelength tuning. Sections (b) and (e) are involved in photon emission.

Figure 3.2 shows the basic architecture of an SG-DBR laser and shows how the five different sections of the laser are aligned. Figure 3.2 is a close-up overhead view of an SG-DBR laser. The basic functionality of the laser can be broken down into two sections: a tuning section and a gain section. The tuning section is involved in selecting the laser's output wavelength, while the gain section is in charge of the emission of photons. The Agility 1107 laser has the same architecture as shown in Figure 3.2.

In Figure 3.2 the SG-DBR laser has three wavelength tuning section sections: front mirror, back mirror, and phase section. An SG-DBR laser's multiple wavelength tuning sections allow it to have a more robust tuning mechanism. Instead of manipulating one variable to reach a particular wavelength, one can manipulate multiple variables. Having multiple wavelength selection variables allows for the implementation of more complicated wavelength selection algorithms. This process will be discussed in more detail later.

Though there are three degrees of freedom when it come to wavelength tuning, the research involved in this report DC biases the phase section of the laser so that only the front and back mirrors are considered in this report. Both the SOA section and gain section are also DC biased. The goal of this report is to explore the feasibility of SG-DBR lasers in OCT therefore keeping these values constant will allow for fewer degrees of freedom. To understand how the laser operates, the lasing conditions will be discussed.

3.3 The Composition of SG-DBR Lasers

The basic functionality of any laser diode is explained by the electron-photon interactions within the device [5]. For a widely tunable laser, the interaction is between the semiconductor's electrons and incident photons. The product of these interactions determines the magnitude, stability, and output wavelength of the laser.

This section will take a quick look at the relation between the carrier density of holes/electrons and a material's index of refraction [5]. This relation is the fundamental tuning mechanism used to control diode lasers [5].

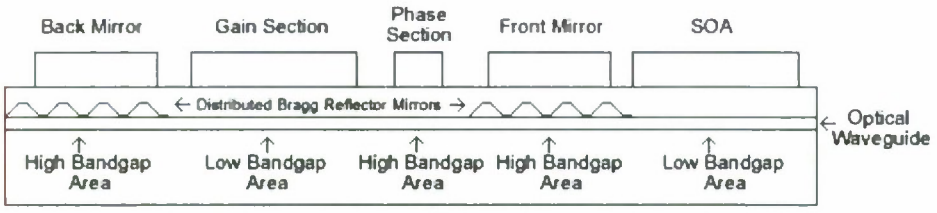


Figure 3.3 - Agility 1107 cross-section. The cross sectional diagram of Agility 1107 CW Widely Tunable Laser Architecture shows how band gaps are distributed throughout the waveguide. Tuning sections have a high band gap while emission/amplification sections have a low band gap.

Figure 3.3 shows the architecture of the SG-DBR laser used in this report. The gain section in Figure 3.3 provides the optical gain of the laser diode while the mirror sections in Figure 3.3 consists of a Bragg grating and is responsible for wavelength selection. Both of these portions of the laser will be discussed in this chapter of the report.

The main functionality of the laser depends on the carrier photon interactions within the device. The band gap differences at each section in the laser allow for stronger/weaker interaction in sections where they are desirable. Figure 3.3 shows the band gaps within different sections of the laser. Low band gaps enable spontaneous emission and the amplification of photons through stimulated emission at the wavelength of interest. This is important for the emission sections of the laser (gain and SOA in Figure 3.3) which need to generate photons. High band gaps are desirable for sections which need to manipulate emitted photons (front mirror, back mirror, and phase in Figure 3.3).

In the discussion of this report, the laser's tuning mechanism will become very

important. In order to develop a tunable ramp for OCT applications, the tuning mechanism for this laser must be explored. The next section explores the Agility 1107's tuning section.

3.3.1 Tuning Section

To explain the SG-DBR laser's tuning mechanism, the mirror sections and the distributed Bragg reflector mirror within the mirror sections must be explained first. The mirror sections of the laser have the same composition as the gain section except the mirror sections have a higher band gap and a distributed Bragg grating embedded within them allowing for wavelength selection. The higher band gap within the mirror sections (shown in Figure 3.3) helps make the mirror sections look like a passive lossless element at the lasing wavelength. Before getting into more detail on the mirror sections, the sampled grating or Distributed Bragg Reflector Mirror (as labeled in Figure 3.3) will be discussed.

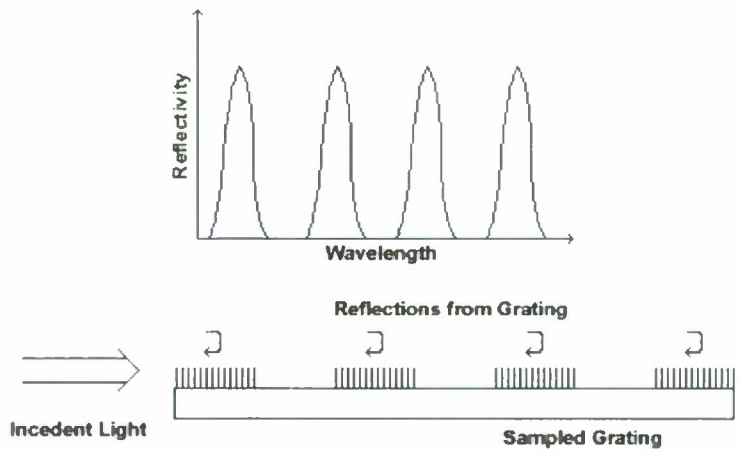


Figure 3.4 - Distributed Bragg Grating. (Left)The gain vs. wavelength function of the mirror section of the laser. (Right) An illustration of the Distributed Bragg reflector grating.

Interference between portions of the Bragg grating causes strong periodic reflections at the peaks shown in the reflectivity versus wavelength plot in Figure 3.4. The Bragg

Reflector Sampled Grating mirror is show on the right in Figure 3.4, its reflectivity diagram is shown to the left in Figure 3.4.

$$\Delta\lambda = \frac{\lambda^2}{2nL} \tag{3.1}$$

(3.1) shows how wavelength periodicity is related to mirror spacing. In (3.1) λ is the lasing wavelength, L is the spacing between mirrors. n is the index of refraction. (3.1) shows that the distance between lasing modes should remain constant. Therefore, a tuning mechanism which slides the reflectivity spectrum is expected. Figure 3.4 shows the constant spacing between reflectivity peaks.

The mirror section is a PIN (p-type, intrinsic, n-type) diode. When a mirror section is biased, the injection current shifts the reflectivity peaks seen in Figure 3.4 (a) by changing the carrier density inside the mirror [6]. This leads to the possibility of mode/wavelength tuning. Each portion of the grating varies the refractive index slightly. The spacing between peaks is dependent on the periodicity of the sampled grating [5]. For the Agility 1107 the spacing between peaks is 5 nm.

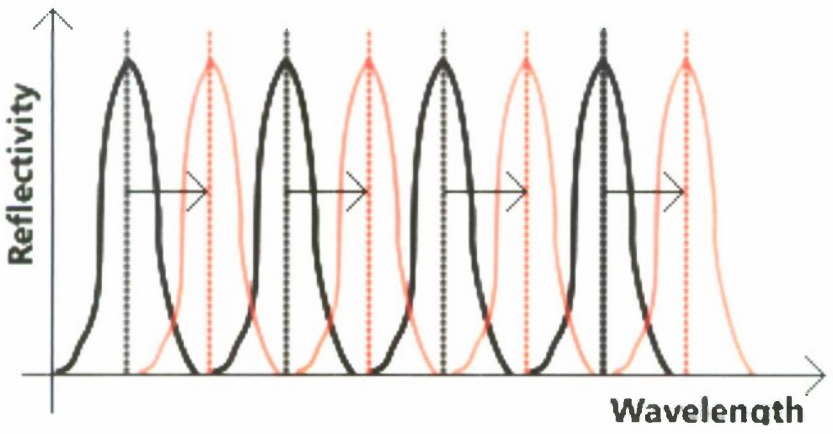


Figure 3.5 - Mirror tuning illustration. The reflectivity vs. wavelength function of the front mirror section of the laser can be altered by changing the carrier density in the mirror. The dark set of peaks have been set

to an initial front mirror current setting and the light set of peaks show how the reflectivity peaks move as current is injected into the mirror.

Figure 3.5 shows the shift in reflectivity spectrum as the injection current changes. The dark spectrum seen in Figure 3.5 is the initial reflectivity spectrum, the lighter spectrum in Figure 3.5 is the new reflectivity spectrum. The arrow in Figure 3.5 shows the sliding direction of the reflectivity.

3.3.2 Lasing Conditions

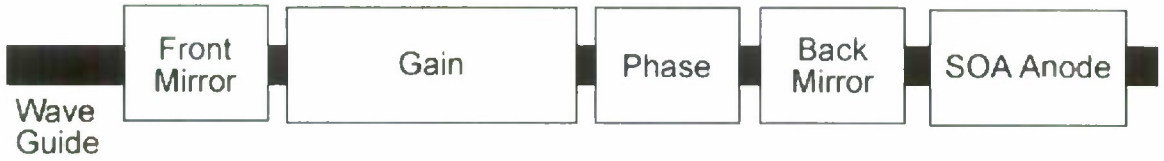


Figure 3.6 - Agility 1107 CW Widely Tunable Laser Architecture.

Figure 3.6 shows the architecture of the Agility 1107 laser. The lasing condition for the laser involves two conditions:

Roundtrip gain (K) > 1 (condition - 1)

Roundtrip Phase ($\varphi = 2\pi n$) (condition - 2)

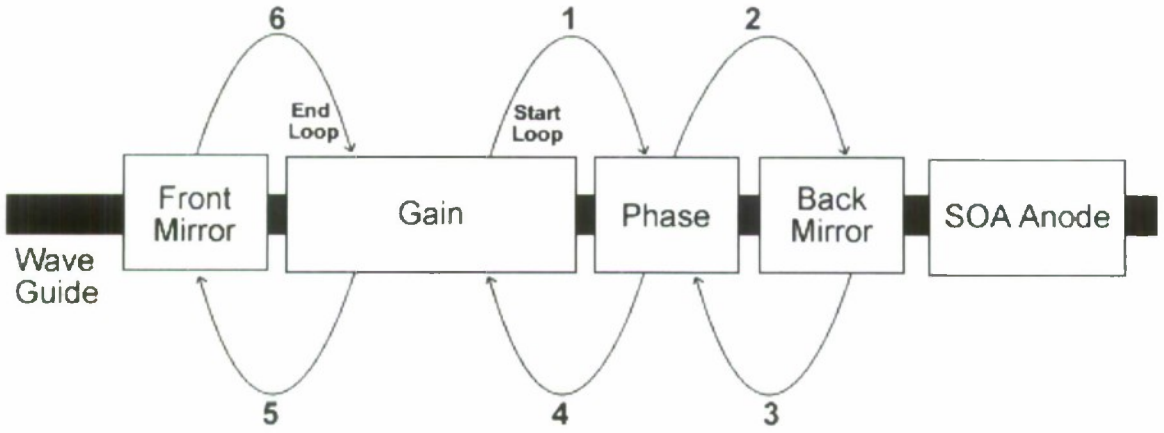


Figure 3.7 - Lasing condition diagram. Agility 1107 CW Widely Tunable Laser Architecture.

The laser is essentially an oscillator. Its function can be thought of as analogous to a Fabry-Perot Interferometer. Basically, the gain section in Figure 3.7 emits photons which are reflected between the two mirrors. The mirrors act as a band pass filter only allowing specific wavelengths to be emitted form the laser. Figure 3.7 shows the path the light travels. By following this path and making a round-trip back to the gain section an equation for lasing condition can be constructed. Each number in Figure 3.4 represents a contribution from a portion of a laser.

$K_{\#}$ = Gain Contribution of last on path #

K_1 = Gain

K_2 = Phase = 1

K_3 = Reflectivity of Back Mirror as a function of Wavelength (frequency dependent)

K_4 = Phase = 1

K_5 = Gain

K_6 = Reflectivity of Front Mirror as a function of Wavelength (frequency dependent)

Oscillation condition on gain:

$$K_1 K_2 K_3 K_4 K_5 K_6 > 1 \qquad (3.2)$$

Therefore a wavelength with roundtrip gain greater than one will cause oscillation. If the laser oscillates, any wavelength in Figure 3.4 with a gain less than one will drive the output to zero in accordance to (3.2).

Figure 3.8 shows the output spectrum of the Agility 1107 laser. The principles behind the shape of the spectrum seen in Figure 3.8 will be discussed along with the tuning

mechanism. Basically, the reflectivity of the mirror sections strengthens the roundtrip gain at particular wavelengths giving those wavelengths a lower loss. The lasing peak in Figure 3.5 occurs at the wavelength where both a peak in the reflectivity of the front mirror and back mirror directly overlap in wavelength.

The tuning mechanism of a generic tunable laser works like a Fabry-Perot interferometer [8]. As a Fabry-Perot interferometer, two mirror elements surrounding the gain section can be adjusted to select a frequency [8]. However in the case of tunable lasers, rather than physically moving the mirrors, the mirror's reflectivity is altered by an injection current.

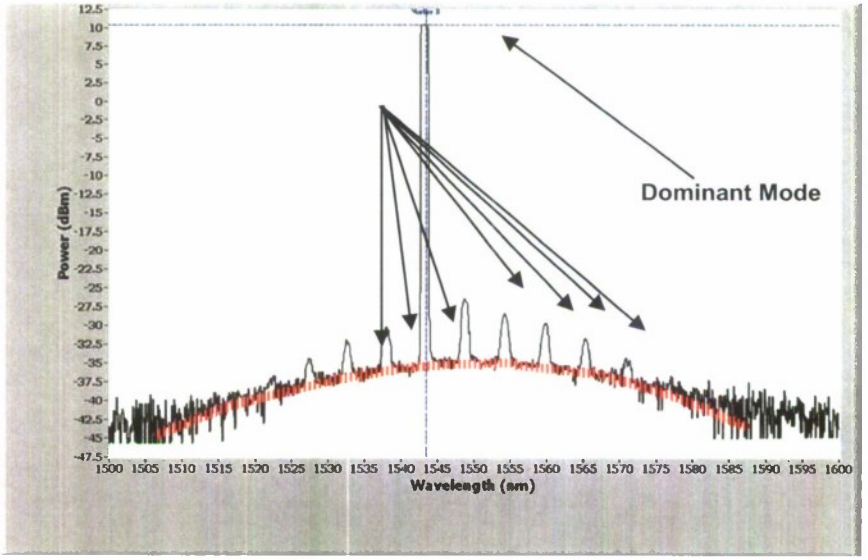


Figure 3.8 - Agility 1107 output spectrum. Laser Output with the dominant mode which satisfies the lasing condition shown.

The modes of the Agility 1107 are signified by the arrows in Figure 3.8. The dominant mode in Figure 3.8 shows a dominant wavelength in which the roundtrip gain is strongest. The spacing between mode peaks is 5nm, the reason for this

spacing is due to the separation distance between Bragg mirrors.

3.4 Wavelength Tuning

The laser is tuned by multiplying front and back mirror gains in condition - 1 (shown in section 3.3.2). The dominating mode seen in Figure 3.9 can move to one of the other non-dominant mode locations. This tuning characteristic is a very important part of widely tunable lasers and will be discussed in detail in this section.

The two mirrors work together as a band pass filter to pass particular wavelengths. The front mirror and back mirror have particular reflectivity's which can be manipulated to allow emission of certain wavelengths. Figure 3.9 illustrates the relationships between the front and back mirror during the wavelength selection process.

Figure 3.9 shows that if there is a strong roundtrip gain for a particular wavelength for both the front and back mirror, the laser will lase at that particular wavelength. This alignment of strong roundtrip gain in Figure 3.9 satisfies (3.2).

Figure 3.9 shows that the reflectivites of the mirrors can add either with strong roundtrip gain or poor roundtrip gain. If they form a roundtrip gain as a function of wavelength that is greater than one, the laser will lase at a particular wavelength. The two mirrors work in conjunction to act as an optical band pass filter.

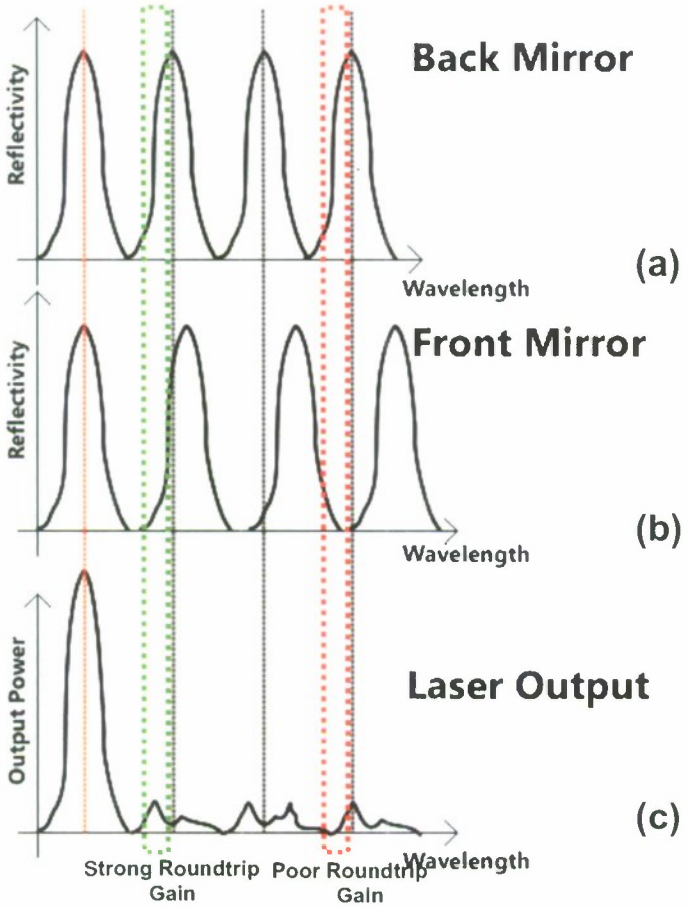


Figure 3.9 - Wavelength tuning illustration. (a) Illustration of the back mirror reflectivity peaks. (b) Illustration of the front mirror reflectivity peaks. (c) Illustration of the laser output due to the alignment of front and back mirror reflectivity peaks.

In Figure 3.9 the reflectivity peaks of the front and back mirror are designed to be unequally spaced by changing the sample period of the sampled Bragg grating [6]. This allows for only one set of peaks to be aligned at a time and allows for the mode/wavelength tuning shown in Figure 3.9.

Figure 3.9 also displays the effect between the two mirrors called the Vernier Effect. By changing the current in the front and back mirror one can shift the reflectivity

peaks of the mirror, this is shown in Figure 3.9. The main idea between using two mirrors is that the Vernier Effect between the two mirrors allows them to tune beyond the range which their own refractive indexes would allow them to tune [6].

Figure 3.5 showed how changing the injection current changes the mirrors peak locations. An increase in injection current raises the electron density in the mirror, therefore changing the index of refraction and causing a shift in the reflectivity spectrum of the mirror. Figure 3.10 shows the 5nm spacing between modes.

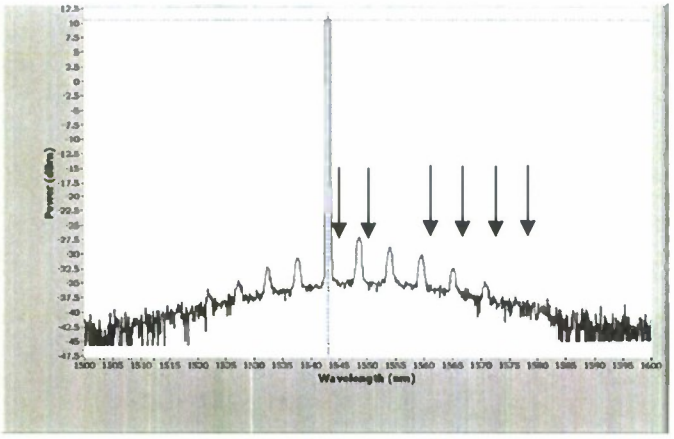
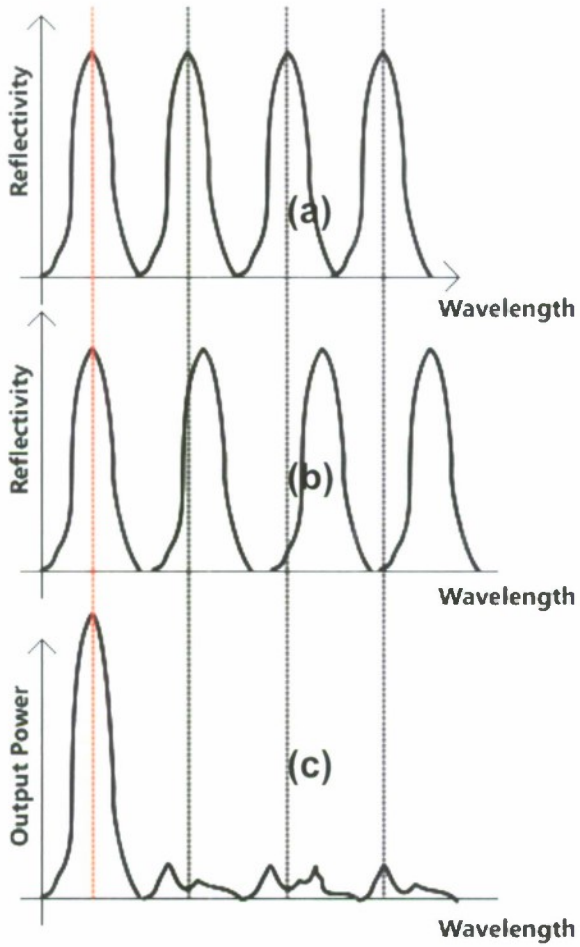


Figure 3.10 - Illustration of 5nm spacing between modes. The diagram shows the 5nm spacing between modes which can be found in (3.2).

This slide in the reflectivity versus wavelength is not only caused by current but any phenomena which can change the carrier density in the front and back mirror. Temperature change of the laser can also change the carrier density within the mirror sections and therefore change the reflectivity versus wavelength spectrum of the mirrors.


Reflectivity
(a)
Wavelength
Reflectivity
(b)
Wavelength
Output Power
(c)
Wavelength

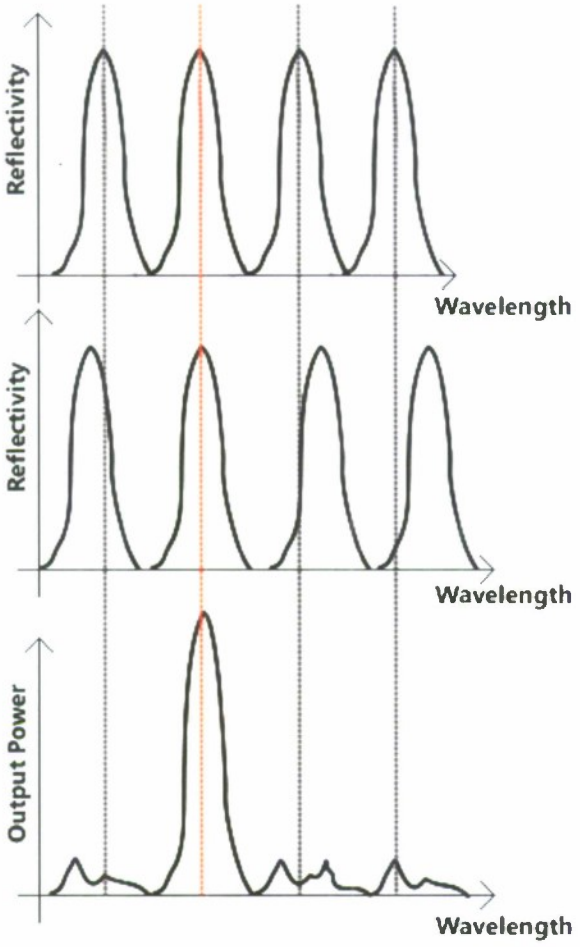


Figure 3.11 - Detailed wavelength tuning using the Agility 1107. (a) The left figure shows the initial back mirror reflectivity while the figure to the right shows the altered back mirror reflectivity. (b) Both the right and left figures show the front mirror reflectivity kept constant during the experiment. (c) An experiment illustrating of how the laser's output changes as the mirrors are tuned. The left figure shows the initial laser output while the right figure shows the laser output offset to the right. For the Agility 1107 this offset distance is 5nm.

When there is an overlap in peaks in the reflectivity spectrums of both mirrors the lasing frequency will change to the frequency of overlapping peaks. This is the basic concept behind the tuning mechanism. Figure 3.11 shows how the output wavelength changes as current in one the mirror's current is varied. This allows the

wavelength to hop from mode to mode. Notice the mode spacing is constant. Each time the laser switches modes, it traverses the same distance (5nm). This is due to the construct of the distributed Bragg grating embedded in the mirror sections of the laser (introduced in Figure 3.4). Both reflectivity spectrums are constant, meaning they both have constant wavelength spacing within their reflectivity spectrums.

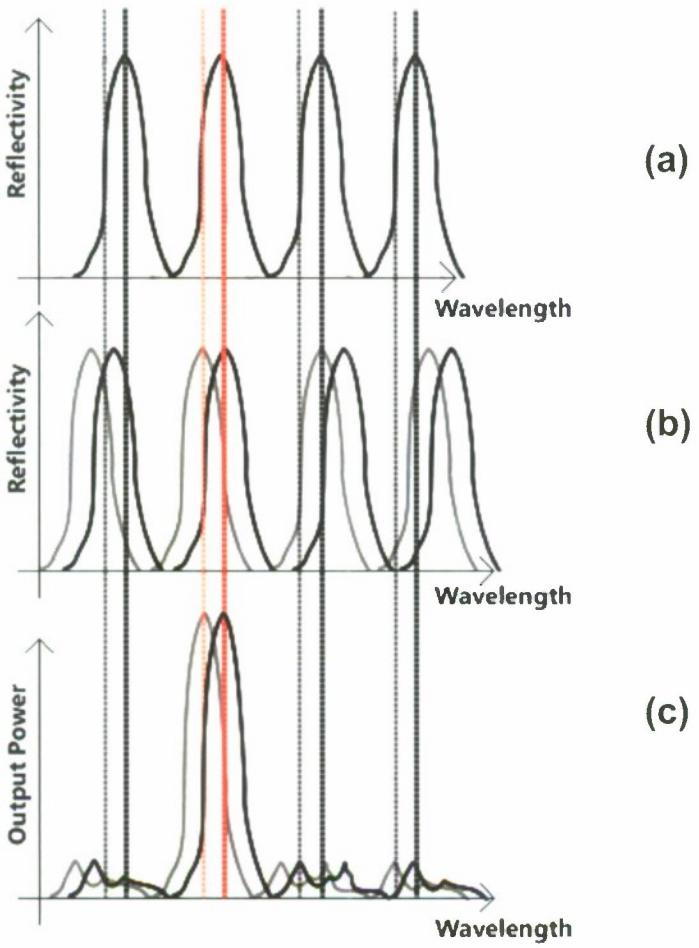


Figure 3.12 - Mode sliding with the Agility 1107. (a) An illustration of back mirror current density being altered slightly and shifting the reflectivity spectrum. (b) An illustration of front mirror current density being altered slightly and shifting the reflectivity spectrum. (c) If both the front and back mirror current densities are altered simultaneously, and mode slide will occur at the output.

Figure 3.12 shows that not only does the architecture of SG-DBR lasers allow for mode hopping, it also allows for mode sliding. By changing the injection of both mirrors simultaneously a mode can slide through a set of wavelengths before switching modes. This characteristic of mode sliding is very important in the construction of a tuning ramp for OCT applications. This sliding aspect allows one to achieve wavelengths one would normally hop over when keeping a mirror constant.

3.5 Feedback Control/Closed Loop Tuning

In most telecommunications applications the SG-DBR laser is aided by a feedback signal that is used for frequency stabilization. This signal is generated using a reference optical filter [9]. The purpose of the feedback loop is to control wavelength drift and protect against inherent wavelength changing attributes such as temperature change and hysteresis. The disadvantage closed loop tuning has when compared to open loop tuning is that it slows down the tuning processes. Closed loop tuning requires constant wavelength verification done through electrical feedback, and therefore is a lot slower than the nanosecond switching capabilities of an open-loop system.

However, the benefit of closed-loop tuning is that less analysis of the laser's characteristics needs to be done upfront, since a feedback mechanism will increase or decrease the injection currents until the correct operating point is reached.

3.5.1 Open Loop Tuning

For fast wavelength switching without on-the-fly diagnostics, wavelength open-loop tuning can be used. In order to use open loop tuning a lookup table must be generated to guide the tuning processes. The lookup table will describe the way the

currents must be manipulated in order to achieve the desired laser output, in the case of OCT a tuning ramp.

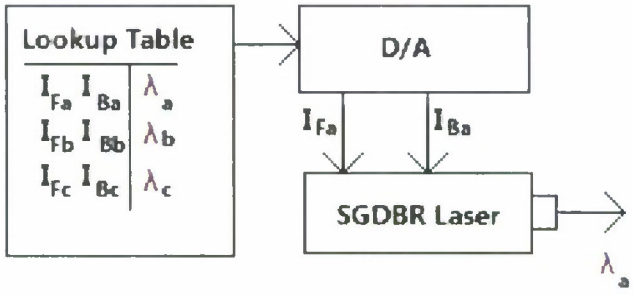


Figure 3.13 - Tuning procedure for the Agility 1107.

The positive aspect of open loop tuning is that the limit to the tuning becomes the D/A converter controlling the injection currents into the laser. This can allow for nanosecond wavelength tuning. When using open loop tuning all the data analysis is done upfront (tuning characteristics, hysteresis, temperature dependence) and taken care of before the lookup table is made.

The negative aspect with open-loop tuning is that one needs a separate lookup table for different operating conditions such as different operating temperatures.

The main issue is that a lookup table for the tuning ramp is designed with a particular current density expectation for each wavelength. A temperature difference varies the electron density, changes the index refraction of a mirror, and essentially emits a different wavelength, changing the operating point of the laser. However using a thermoelectric cooler (TEC) will help eliminate this issue.

Another problem an open-loop tuning mechanism must overcome is the inherent

hysteresis associated with wavelength switching in SGDBR lasers. An open-loop tuning algorithm must avoid hysteretic regions of the laser. This will involve hopping over hysteretic regions to different stable modes at a particular wavelength.

Sampled grating lasers overcome the limitations of narrowly-tunable DBR lasers [10]. Work has been done in creating a closed look tuning ramp or using on external lab instruments [10]. In this application the lasers functionality in an open-loop environment is explored.

3.6 Application: Telecommunications

Currently, in near-term applications (communications) there is no demand for fast wavelength switching [7]. This report is looking to explore the imaging application of SGDBR laser where the fast switching potential is not only useful, but mandatory for real time applications.

However, in communication applications the fast wavelength switching capability of DBR lasers is not used or desired. A the fast wavelength switching capabilities of DBR lasers can cause unwanted wavelength change during a data transmission due to the fact that data needs to be transmitted at a particular frequency. Therefore any noise in the current source powering the laser could cause an unwanted frequency change.

(a)

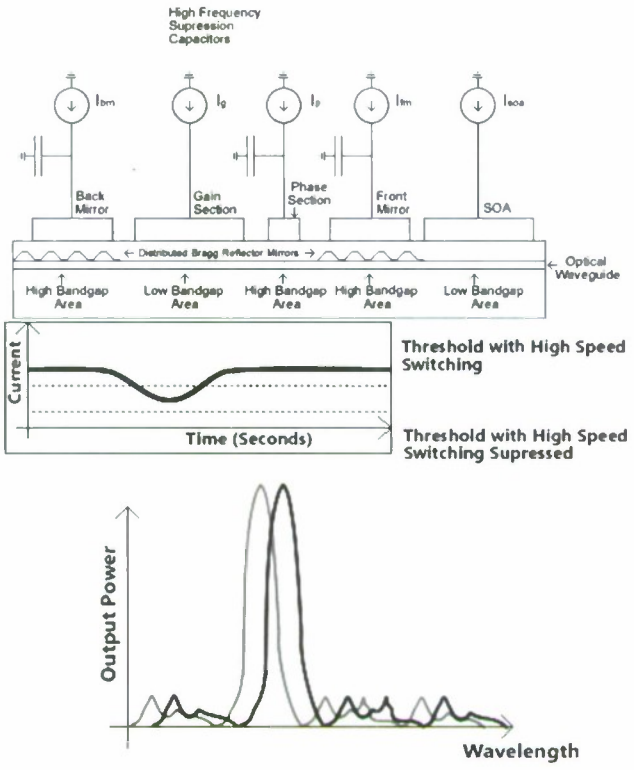


Figure 3.14 - A teleeommunications protection feature of the Agility 1107. (a) Placing capacitors at the tuning section input create a low pass filter and prevent unwanted wavelength change. (b) An illustration showing how a small perturbation in the signal has the potential to change the Agility 1107 output.

This high speed switching is suppressed using capacitors. The capacitors act as a low pass filter upon the injection currents which don't allow them to change rapidly and in turn effect the laser's output. This allows for protection against frequency change due to sudden noise. Figure 3.14 (a) shows how adding the capacitors change the switching current threshold. Figure 3.14 (b) shows the effect sudden noise could have on lasing frequency. Without this threshold, the frequency of the laser would have changed.

Any phenomenon which has an influence over the current density of a laser can

affect the lasers operating point. For example, thermal tuning caused by electrical power dissipation during a sudden current change has the ability to change a lasers frequency by 20Ghz[6]. This is important in wavelength division multiplexing.

3.7 Overview of Physics Concepts

This section may be skipped without loss of understanding of how the Agility 1107 is used in the application presented in this report.

3.7.1 Gain Section and Spontaneous/Stimulated Emission

The gain section of the laser is an optical amplifier whose function can be understood by the theories of stimulated and spontaneous emission. Stimulated emission in a tunable laser consists of photon motion perturbing electrons, causing the electron to be demoted into the valence band; allowing a companion photon to escape in the process. The companion photon has the same frequency, phase, direction of motion, and polarization as the incident photon [15].

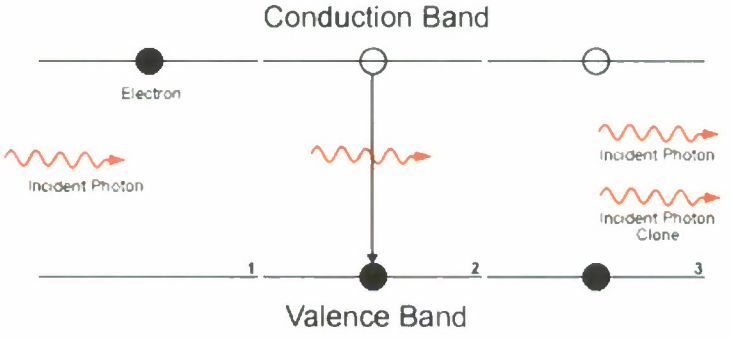


Figure 3.15 - An illustration of stimulated emission.

Figure 3.15 show the stimulated emission effect. A photon in Figure 3.15 passes through the conduction band and valence band, allowing an electron to move into the valence band, the process leads to the duplication of a photon. On the other hand spontaneous emission is the emission of light due to carrier transition between

bands not involving incident photons.

CHAPTER 4 Tuning Characteristics

The goal of this section is to produce an algorithm for the Agility 1107 CW Widely Tunable Laser (Agility 1107) which will ensure a reproducible and reliable wavelength versus time ramp output to be used in an OCT architecture. This section will explore the procedure, feasibility, and reliability of creating a tuning ramp of wavelength versus time using the Agility 1107.

This chapter will start by proving feasibility of producing a tuning ramp via the Agility 1107. The discussion will start by taking the reader through the multiple experiments performed on the Agility 1107 to explore its tuning characteristics. It will then go into an analysis of the characterization of the Agility 1107. Finally, the discussion will then go into some problems which will need to be avoided in order to produce a reliable tuning ramp.

With respect to lasing wavelength, the laser has three degrees of freedom: front mirror current, back mirror current, and phase current. This report chooses to fix bias the phase tuning abilities of the laser and concentrates on characterizing the front and back mirror effects on the lasing wavelength.

4.1 Feasibility of Wavelength Tuning

The first experiment performed using the laser was to see if there even was a possibility of creating a wavelength ramp without gaps in wavelength coverage using the Agility 1107. The experiment involved recording a sample of every possible operating point for the laser and ordering those points in terms of resulting wavelength.

4.1.1 Feasibility of Wavelength Tuning Experiment Setup

The experiment setup involved interfacing hardware to the laser in order for automated operation and measurement control. This automation was done using the Labview programming environment. The Labview Virtual Instrument (VI) for the experiment described in this section can be seen in Appendix B.

Figure 4.1 shows the hardware configuration used in the tuning ramp feasibility experiment. Basically, the Labview Interface in Figure 4.1 sends voltage values to the D/A (Labjack U12) via USB. Those voltages are then converted via a V-I converter to input currents for the laser. The thermoelectric cooler (TEC) controller in Figure 4.1 was used to maintain temperature stability during the experiment. The Optical Spectrum Analyzer in the experiment was used to collect the laser's output wavelength data.

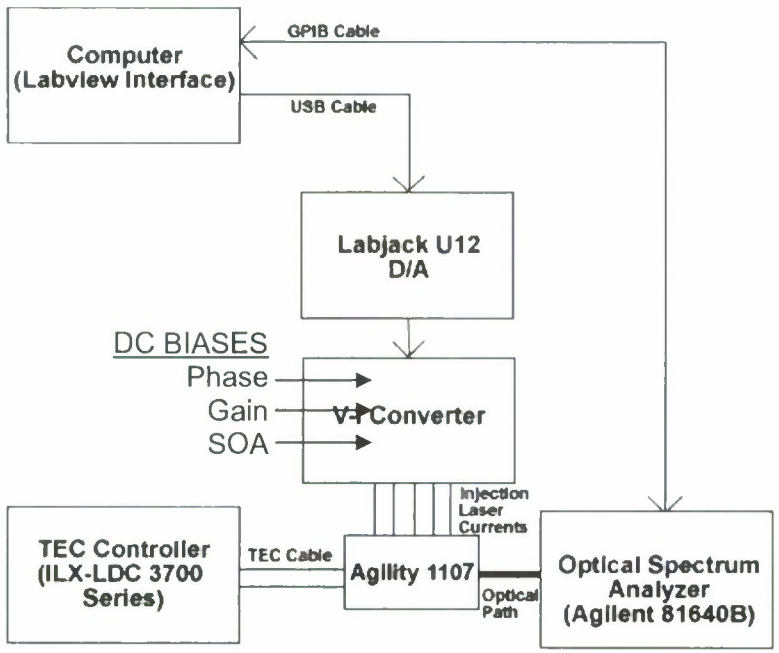


Figure 4.1 - Experimental setup for the ramped tuning feasibility experiment. The Labjack U12 D/A

controls the interface board which is used to send currents into the front and back mirror sections of the laser.

At times in this report, the laser's input variable for the front and back mirrors will be volts due to the nature of the experiments and use of a VI converter to take measurements. The current conversion factor for laser wavelength data plotted with respect to voltage are show in (4.1) (back mirror conversion factor) and (4.2) (front mirror conversion factor).

$$i_{BackMirror}(v) \approx \frac{v-1.4}{60} \qquad v \geq 1.4 \tag{4.1}$$

$$i_{FrontMirror}(v) \approx \frac{v-1.4}{110} \qquad v \geq 1.4 \tag{4.2}$$

4.1.2 Feasibility of Wavelength Tuning Experiment Definition

The goal of the experiment is to characterize the Agility 1107 Laser. More specifically, the objective is to map the input variables of the laser control system shown in Figure 4.1 to the output variables of the system. The system variables are shown in Figure 4.2. The input for the system in Figure 4.2 will be front and back mirror voltage while the output variable will be lasing wavelength.

Figure 4.2 - Diagram of laser control system. Experimental setup for the ramped tuning feasibility experiment.

The experiment used to characterize the Agility 1107 was defined like so: While the front mirror is held constant, the back mirror voltage is ramped from 1.4V (0 ma) to

5V (60ma) in 5mV (.83mA) increments. When the back mirror voltage reaches 5V (60mA) the front mirror's voltage is also incremented 5mV (.45ma) between voltage values of 1.4V (0ma) to 5V (32ma). A diagram of the experiment is shown in Figure 4.3.

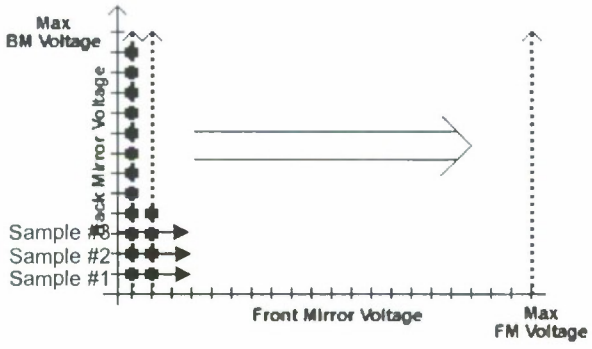


Figure 4.3 - Diagram of tuning ramp feasibility experiment/Agility 1107 laser characterization experiment. An input voltage upon the back mirror was ramped for every front mirror value.

Therefore, the procedure described earlier and diagramed in Figure 4.2 yields a total of 5184 current combinations. Of those combinations, the goal is to have them sorted by output wavelength in such a way that a linear ramp can be constructed.

4.1.3 Feasibility of Wavelength Tuning Experiment Results

After performing the experiment defined in Section 4.1.2, the data in Figure 4.4 was obtained. Using 5184 input combinations (front mirror and back mirror voltages) for Figure 4.2 and mapping each combination with respect to wavelength, Figure 4.4 was obtained. The sample number represents the point in time the data was taken. The key point of Figure 4.4 is that it shows that the inputs of the system can be ordered in such a way as to tune the laser's output wavelength in a linear fashion (shown in Figure 4.5).

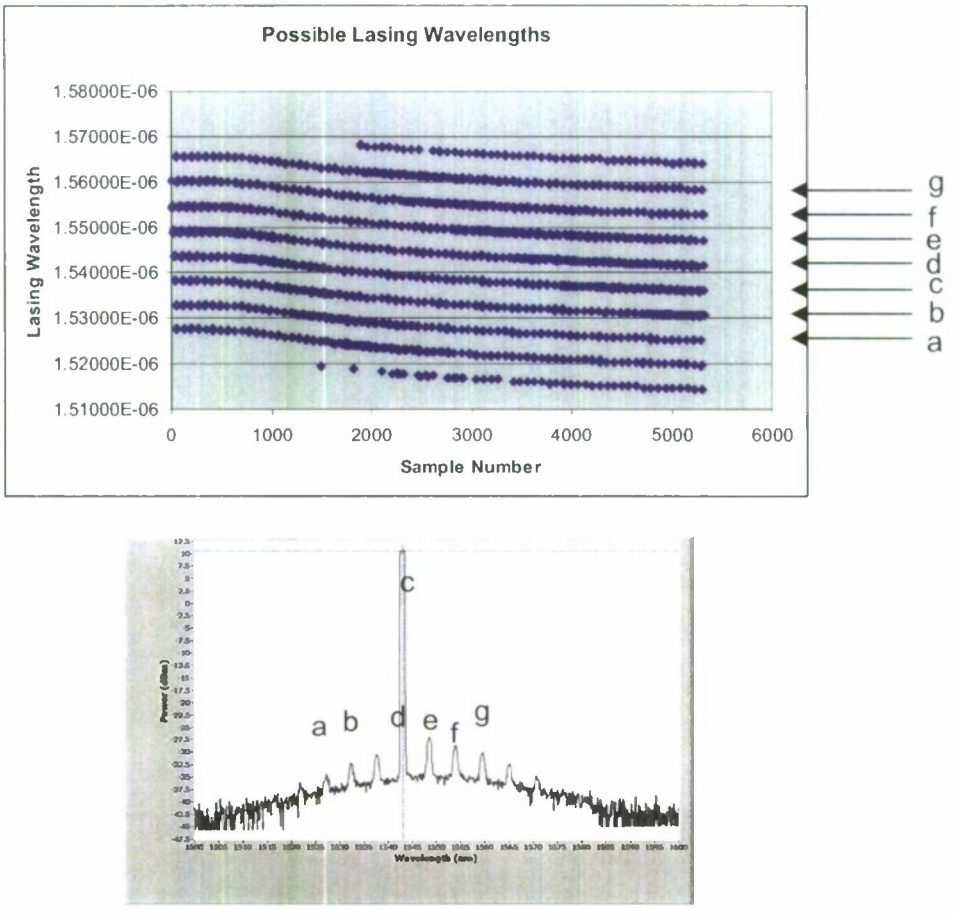


Figure 4.4 - An illustration of mode movement in experimental data. (Top) Diagram of tuning ramp feasibility experiment/Agility 1107 laser characterization experiment. (Bottom) Output spectrum of the Agility 1107 Laser. The letters in the figure label the different modes and relate the upper and lower images to each other.

As seen in Figure 4.4 the lasing range of the laser is from 1515nm to 1568nm. This corresponds to a frequency span from 191 THz to 197 THz (assuming light speed in free space). However, the more stable potion of the graph is from 1520nm to 1560nm. This conclusion was reached due to the fact that the density of points in between 1520nm to 1560nm in Figure 4.4 was large.

Each connected group of points in Figure 4.4 represents a particular mode the laser lases at. The modes are labeled in Figure 4.4 with the letter (a) through (g). Notice in Figure 4.4 that not all modes can reach all wavelengths. For example, in Figure 4.4 mode (a) can lase at 1528nm but will never be able to lase at 1560nm. In this case, 1560nm is not able to be reached by mode (a). In order to lase at 1560nm the lasing mode (g) would need to be used. Specifically, in Figure 4.4 there is a 5nm spacing between each mode and each mode can only slide about 10nm. The aspect that each mode can slide 10nm even though there is only 5nm of spacing between each mode is important in wavelength reliability. If a particular mode is having trouble lasing at a particular frequency, the laser can be manipulated so that another mode which is stable at the particular frequency will lase. For example in Figure 4.4 both modes (c) and (d) are able to lase at 1545nm. This trait comes in to play when dealing with hysteresis.

Again, Using 5184 input combinations (front mirror and back mirror voltages) for Figure 4.2 and mapping each combination with respect to wavelength Figure 4.4 was obtained. Figure 4.5 is a zoomed in version of Figure 4.4. Notice that the points in Figure 4.5 are spaced closely enough and seem abundant enough to form a tuning ramp.

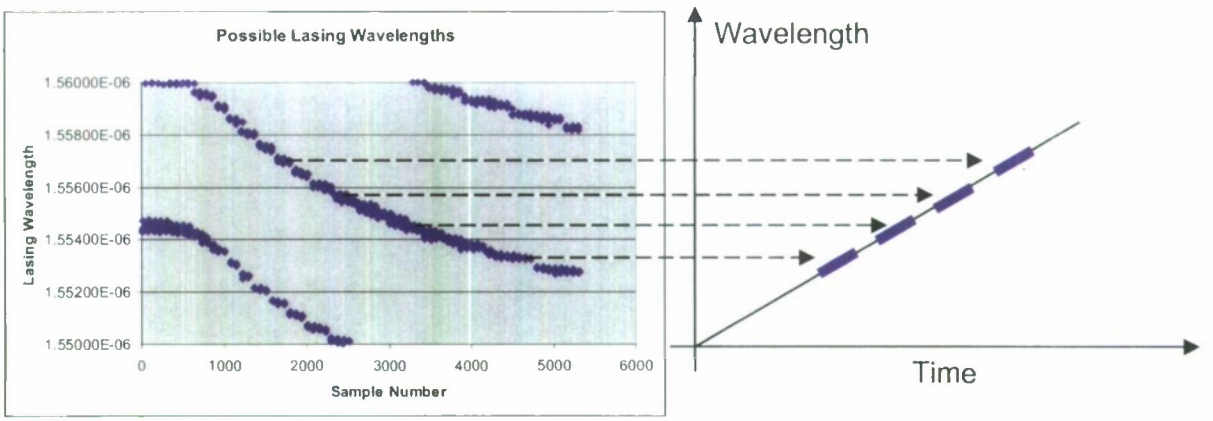


Figure 4.5 - A possible tuning ramp. (Left) A zoomed in version of Figure 4.4 on the vertical wavelength axis. The goal of this illustration is to show that the lasing wavelengths are both close enough and abundant enough to form a frequency tuned ramp from. (Right) A possible tuning ramp.

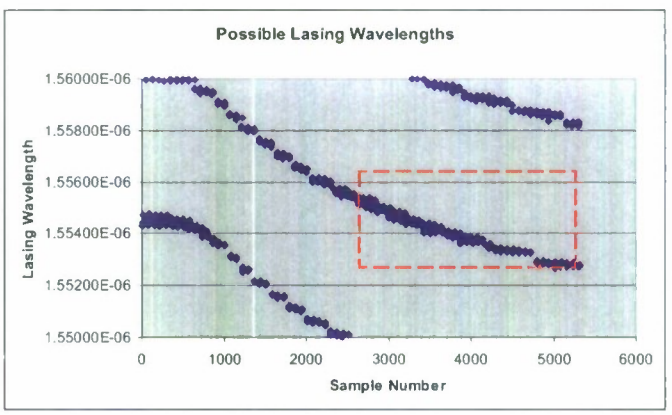


Figure 4.6 - Illustration of a stable lasing region. A zoomed in version of Figure 4.4 on the vertical wavelength axis. The box indicates a large density of points lasing at a particular wavelength between 1540nm and 1550nm.

In Figure 4.6 the region surrounded by the dashed box contains a dense amount of points. The boxed region in Figure 4.6 implies that there is a stable operating region between 1553nm and 1557nm. However Figure 4.4 only shows at what wavelengths the laser can operate. Figure 4.4 says nothing about how to tune the front and back

mirror so that a particular wavelength is reached.

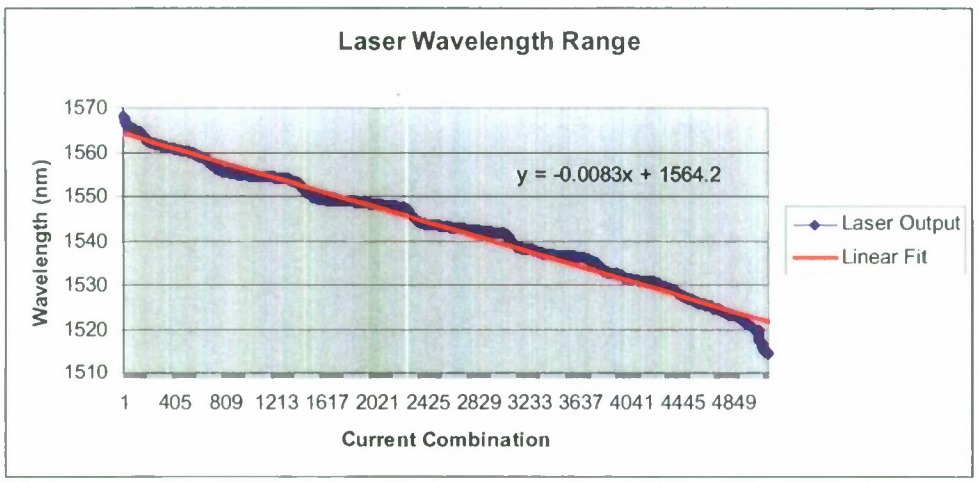
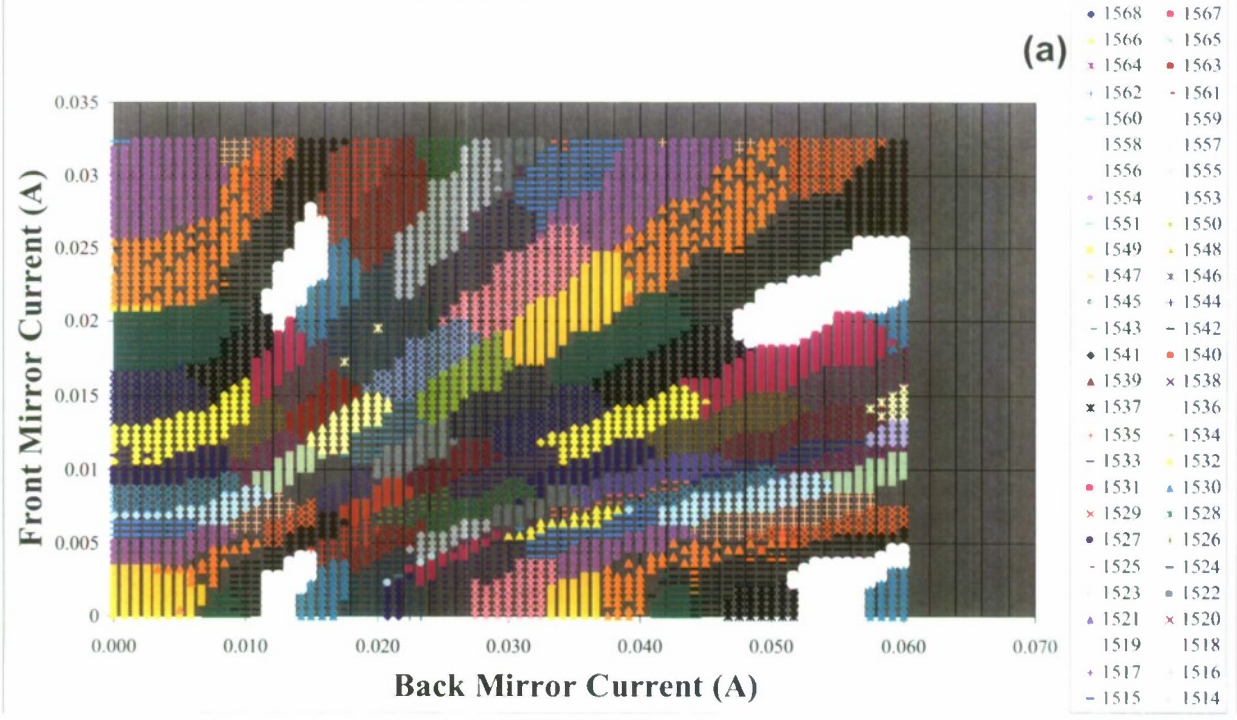


Figure 4.7 - Sorted experimental data with respect to wavelength. The data obtained in the experiment described in Section 4.1.4 sorted with respect to wavelength. The x-axis represents a particular front and back mirror voltage combination input from the experiment described in Section4.1.4.

Using 5184 input combinations (front mirror and back mirror voltages) for Figure 4.2 and mapping each combination with respect to wavelength, Figure 4.4 was obtained. Sorting Figure 4.4 with respect to wavelength yields Figure 4.7. Figure 4.7 shows that if all the current combinations for the front and back mirror are ordered with respect to wavelength, the output spreads into region of smoothness and steep regions. The smooth regions represent the stable operating points while the steep regions represent areas where mode hoping may occur. As a result, the experiment in described in Section 4.1.2 showed that it is feasible to develop a tuning ramp using the Agility 1107 because the system outputs (Figure 4.2) from all the possible system input combinations of Figure 4.2 used in the experiment can be ordered in a fashion such that the output produces a linear ramp.

4.2 Agility 1107 Characterization

This section uses the data from the experiment described in Figure 4.1.2 (seen in Figure 4.4) and plots the data from Experiment 4.1.2 with respect to front and back mirror current. Figure 4.7 directly maps the front and back mirror input current to the output wavelength of the system described in Figure 4.2. The input variables, front mirror and back mirror voltages from Figure 4.4 have been converted to currents (to generalize results). The different colors of each section of Figure 4.8 (a) and (b) represent a different lasing wavelength. Each colored section in Figure 4.8 (a) and (b) describes a wavelength region which is truncated with respect to nanometer wide wavelength bins. For example, 1567.76nm will be found binned under 1567nm.



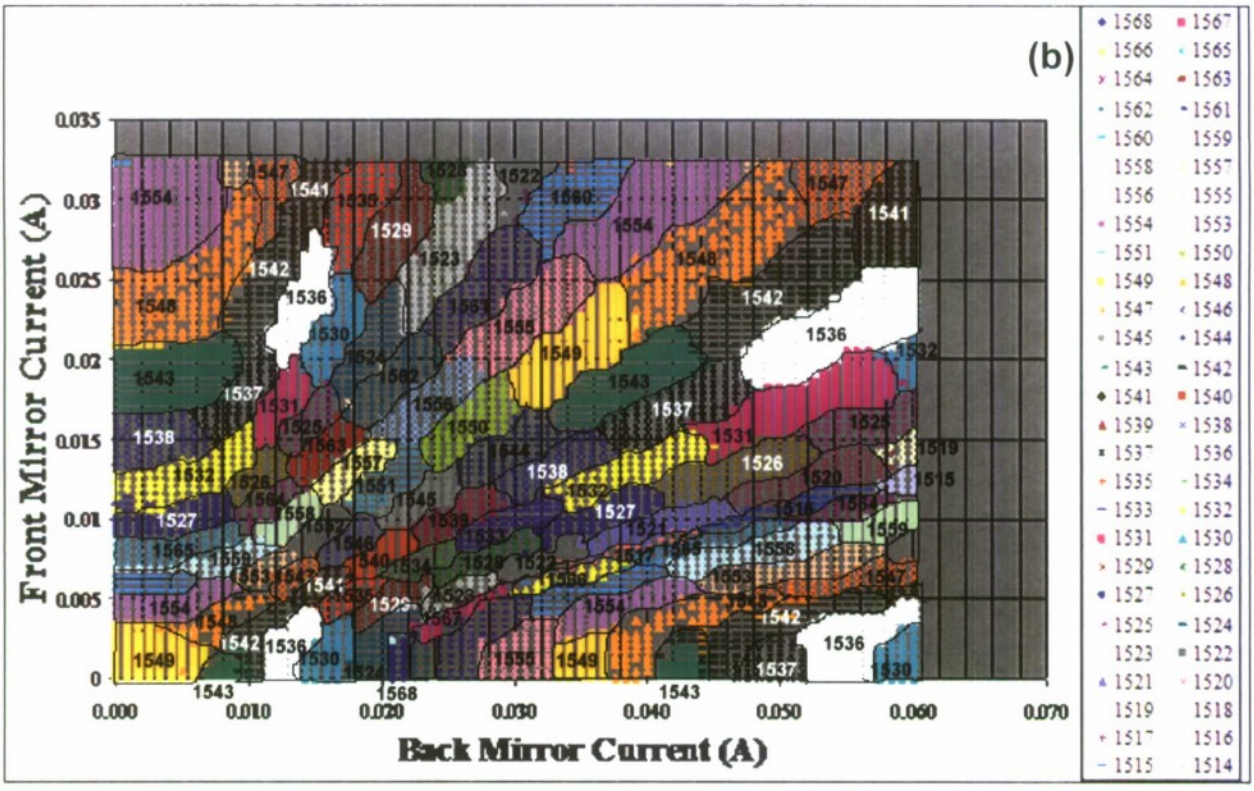


Figure 4.8 - Full characterization of Agility 1107. This illustration shows the results from the experiment described in section 4.1.2. The data is graphed with respect to front and back mirror current, the different colors represent different lasing wavelengths at different front and back mirror current combinations. (a) Non-labeled plot was included for easier viewing of the data. (b) Labeled Plot was included to display data trends.

Figure 4.8 (a) shows that adjacent sections do not always correspond to adjacent steps in wavelength. Therefore a tuning algorithm will need to hop between the different wavelength regions in Figure 4.8 (a) to change wavelength. This idea of hopping from wavelength to wavelength using an open-loop architecture brings with it different issues that need to be dealt with:

1. Is hopping from one wavelength to another repeatable?
2. Does it matter which way one approaches a current operating point (from

lower frequency to higher frequency)?

3. Do thermal characteristics effect the hopping?

Figure 4.8 (a) assumes that the gain, phase, SOA, and thermal characteristics remain constant during laser operation.

Figure 4.9 shows how the modes are spaced within the tuning characteristics. Each solid color region has a 5nm spaced from the color regions directly adjacent to either side of it. Each color in Figure 4.9 represent a modes motion through the tuning characteristics.

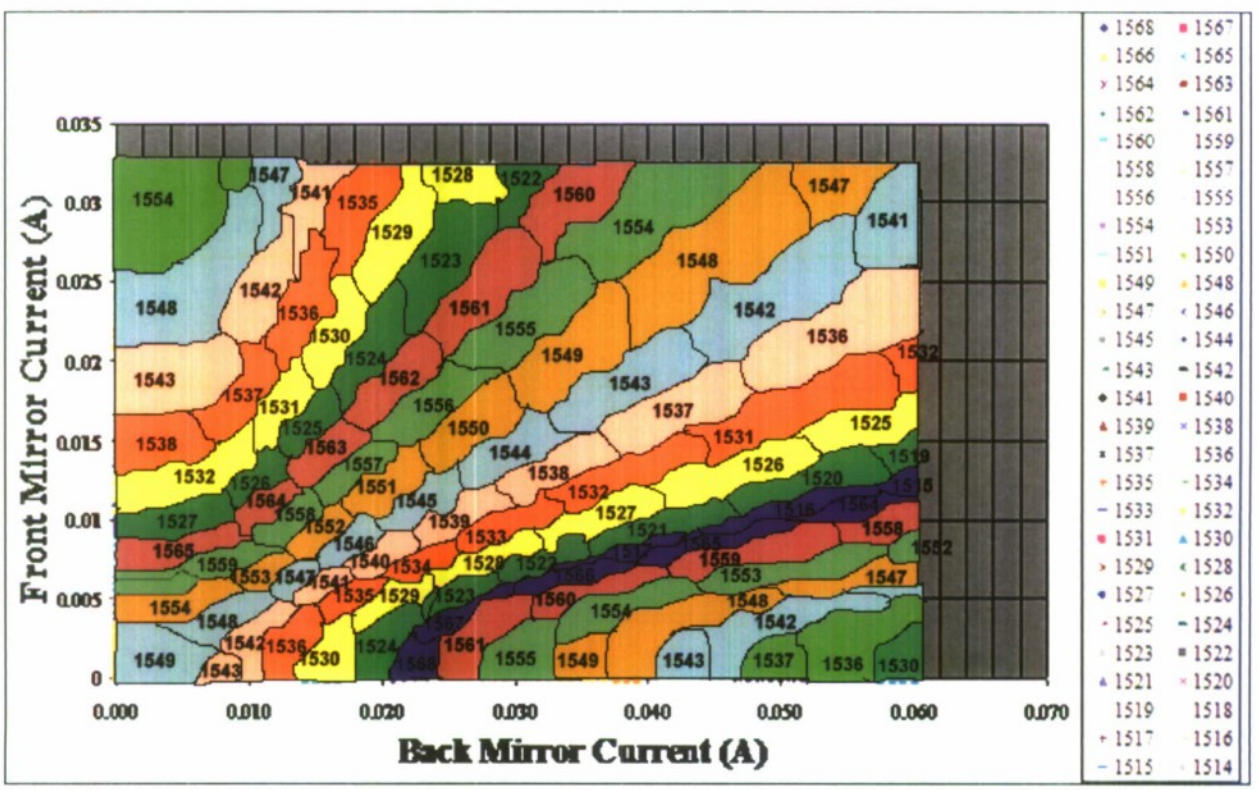


Figure 4.9 - Full characterization of the Agility 1107 showing mode motion. This illustration shows how the laser's modes traverse different sections of the tuning characteristics. Each mode has a particular set of wavelengths for which it can cause the laser to lase.

4.3 Experiment Two: Repeatability Test

The main goal for any open-loop tuning technique is that it needs to be repeatable without feedback. Since an open-loop architecture has no way to reference or check itself, the laser ramp relies highly on an accurate repeatability versus time analysis of the data.

To test for repeatability, the experiment defined in Section 4.2.1 was run multiple times. The results of both experiments were subtracted from each other and Figure 4.9 was the result. In Figure 4.10 the x-axis represents a sample input combination (front and back mirror current values) shown Figure 4.2. The result was 96% repeatability within 5nm, however that assumed that the laser was ramped in a fashion that did not take into account output wavelength. Figure 4.10 shows that one cannot simply run through a wavelength ramp from the lowest frequency to the highest.

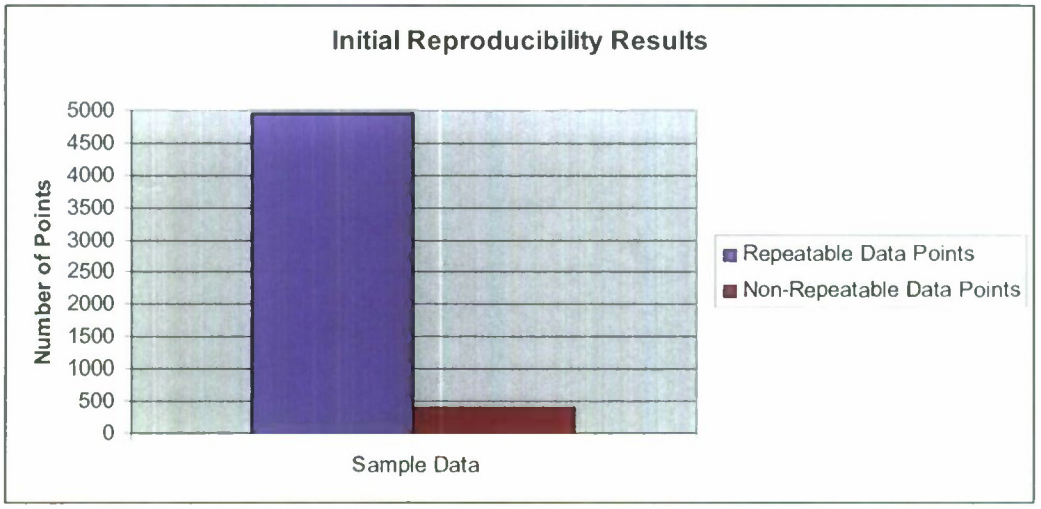


Figure 4.10 - Initial Reproducibility Results.

Figure 4.10 shows the reproducibility results for two initial tuning ramps. It was noticed that when the data was not reproducible the wavelength difference between the two sets of data was 5nm. Note that the wavelength never hops more than 5nm. This means that the worry of hopping into a total different section in unlikely. It was found that large abrupt changes in the injection currents in both mirrors effect the repeatability. This will be dealt with more later.

4.3.1 Experiment 3: Tuning Ramp Repeatability with Respect to Injection Current Ramping (System Repeatability Experiment)

Figure 4.1 shows the hardware configuration used in the System Repeatability Experiment. Basically, the Labview Interface in Figure 4.1 would send voltage values to the D/A (Labjack U12) via USB. Those voltages are then converted via V-I converter to input currents for the laser. The TEC controller in Figure 4.1 was used to maintain temperature stability during the experiment. The Optical Spectrum Analyzer in experiment was used to collect output wavelength data.

4.3.2 System Repeatability Experiment Definition

The repeatability experiment is defined slightly differently than the experiment definition in Section 4.1.2. The System Repeatability Experiment is defined like so: For a constant front mirror voltage the back mirror voltage is ramped from its minimum value to its maximum value (1.4V to 5.0V) in 5mV increments. Then back mirror voltage is ramped down from its maximum value to its minimum value (5.0V to 1.4V) in 5mV increments. Once the back mirror reaches its minimum value the front mirror is incremented by 5mV. This process continues until the front mirror reaches its max value (5V). The process is illustrated in Figure 4.11 (a).

Then for a constant back mirror voltage the front mirror voltage is ramped from its minimum value to its maximum value (1.4V to 5.0V) in 5mV increments. Then front mirror voltage is ramped down from its maximum value to its minimum value (5.0V to 1.4V) in 5mV increments. Once the front mirror reaches its minimum value the back mirror is incremented by 5mV. This process continues until the back mirror reaches its max value (5V). The process is illustrated in Figure 4.11 (b).

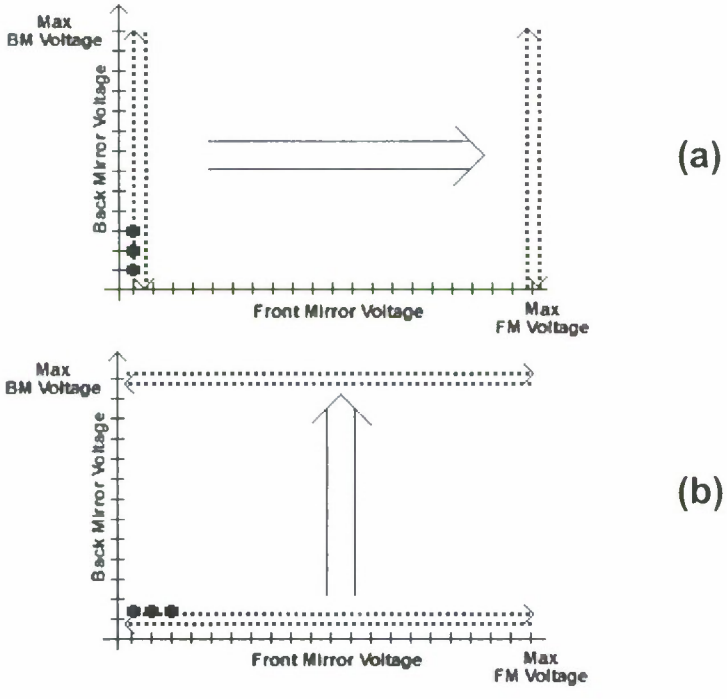


Figure 4.11 - Experiment used to determine repeatable wavelengths. (a) Illustration of the back mirror repeatability test. (b) Illustration of the front mirror repeatability test.

4.3.3 System Repeatability Experiment Results

When performing the repeatability test described in 4.3.3 Figure 4.12 is an example result. Figure 4.12 shows how the output wavelength changes with respect to ramped back mirror current. The main aspect which Figure 4.12 illustrates is that

there seems to not be a 1 to 1 mapping between input front mirror and back mirror voltages and output wavelength for all conditions. Figure 4.12 shows that there are some internal features of the system in Figure 4.2 which do not allow this 1 to 1 mapping. More specifically, after an examination of Figure 4.11, it seems as if hysteresis is occurring. Hysteresis is defined as an input combination to the system in Figure 4.2 which does not yield a unique wavelength output. The following section will attempt to explain this hysteresis.

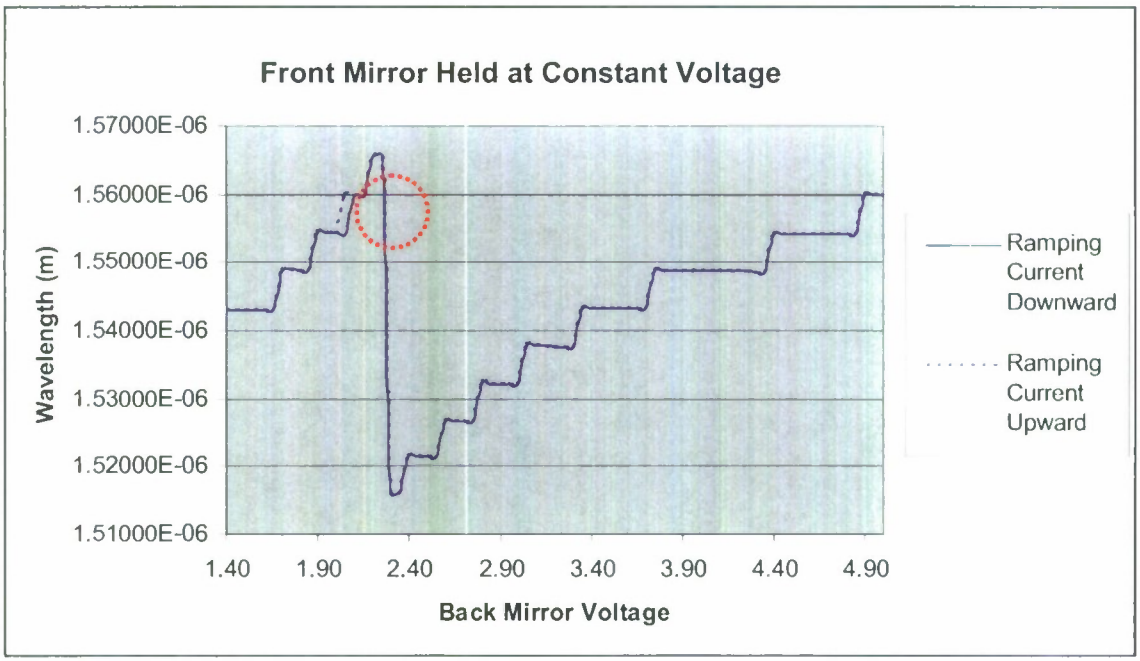


Figure 4.12 - Hysteretic lasing region. An illustration of the repeatability test while holding the front mirror at constant voltage. The hysteretic region is circled.

4.4 Hysteresis

This section will attempt to explain the hysteresis encountered in section 4.3.3.

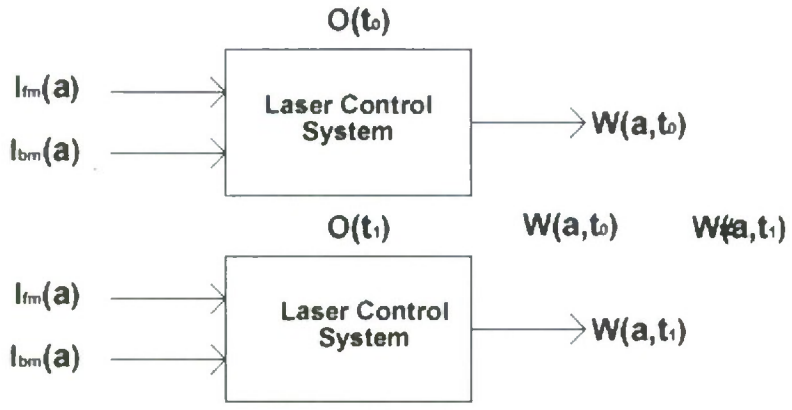


Figure 4.13 - A definition of hysteresis. The same inputs do not always yield the same wavelength output at two different points in time. For example, let a be a voltage value from the D/A which sends a current to the laser's mirrors. Hystesis implies that for the same voltage value at different times, the laser may not lase at the same wavelength and there is a 5nm difference between $W(a,t_0)$ and $W(a,t_1)$

Figure 4.13 shows the definition of hysteresis used in this report. I_{bm} and I_{fm} are input values to the system. O(t) is some function that describes the internals of the system. W(a,t) is a function which describes the output of the system. Hysteresis occurs when the inputs of the system do not yield the same output at two different points in time.

Hysteresis occurs because the Bragg reflectors interact with the light entering the laser through unintended absorption. The resulting absorption changes the carrier density in the mirror section which in turn changes the mirrors index of refraction [5]. Figure 4.14 show that it is energetically favorable for the laser to maintain lasing at its current wavelength before switching. The type of hysteresis observed is a "lag" in wavelength shift. It was deduced that there must be a feedback mechanism inherent in the electronics between lasing wavelength and carrier density.

4.4.1 Signs of Hysteresis

Figure 4.14 shows that for most values of back mirror voltages the wavelengths are repeated as one ramps upward and downward through voltages. However, in some regions wavelength repeatability is not achieved. In Figure 4.14 there is not consistent wavelength repeatability at the two points surrounded by the two dotted boxes. This region is called a hysteretic region.

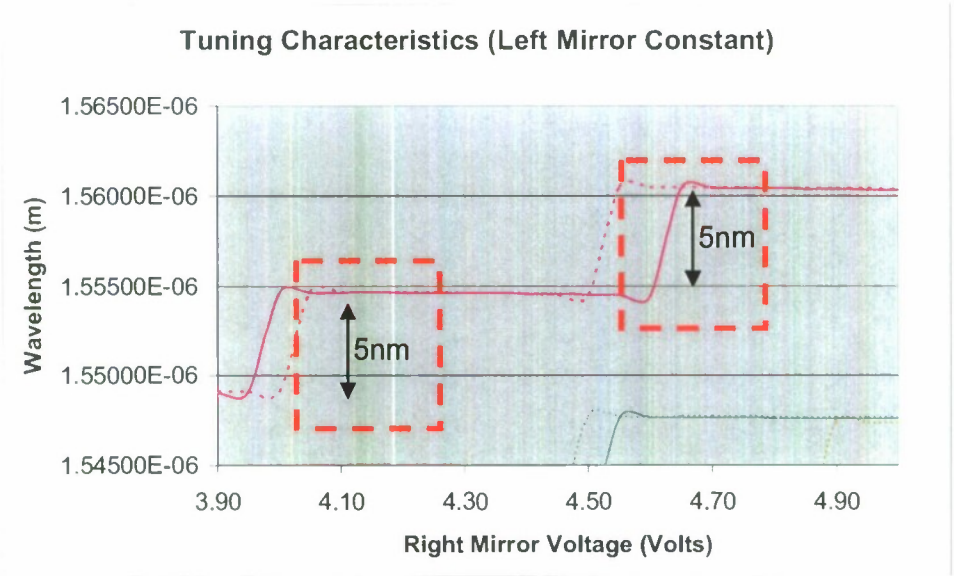


Figure 4.14 - A y-axis close-up of the hysteretic region. An observation that hysteresis always occurred at output wavelengths 5nm from the desired wavelength.

The key aspect of Figure 4.14 is that during hysteresis the difference between the wavelength operating point is 5nm. The 5nm distance also happens to be the spacing between modes. This means when the laser is operating inside a hysteretic region, its exact lasing mode cannot be determined without knowledge of history. Therefore in the design of an open-loop tuning ramp algorithm it would be prudent to avoid these regions. The ways these regions can be avoided is to jump on a more

stable path once a hysteretic region is encountered

4.4.2 Hysteresis Explanation

The phenomena of hysteresis in the Agility 1107 is one where the current lasing wavelength influences electron density through optical absorption in the front and back mirrors. The mirror's band gap is larger than the band gap of the gain section. The goal of designing the mirrors with higher band gaps is to make them look transparent to the rest of the system (since they are only used for wavelength selection). However, absorption still occurs and is the main cause of the hystesis. Optical absorption promotes electrons within the mirrors causes a promotion of electrons into the conduction band, therefore changing the carrier density of the mirror and the refractive index.

4.4.3 Treating Hysteresis

Hysteresis was treated by finding non-repeatable points using the experiment defined in Section 4.3.2. If there was not a 1 to 1 correspondence of input operating points with respect to wavelength output, the operating point was classified as being prone to hysteresis and discarded.

A computer program was used to characterize the laser, spot the hysteresis, and come up with a tuning dictionary for the Agility 1107. This tuning dictionary contained all the stable lasing points where hysteresis did not occur. The process of finding the hysteretic regions is shown in Figure 4.15. The hysteretic regions where found by holding the front mirror constant and scrolling the back mirror between its maximum and minimum values. The two sets of mirror current and wavelength values were stored. One set for scrolling the mirror current upward, and one set for scrolling the

mirror current downward. The program then compared the lasing wavelength for a particular current during an upward scroll and downward scroll. The algorithm then looked for 5nm mode jumps. If a 5nm difference occurred between the upward and downward scroll the operating point is marked as prone to hysteresis and not stored in the tuning dictionary. The tuning dictionary is the file which contains all the stable operating points. When constructing a tuning ramp, the tuning program will chose operating points from that tuning dictionary.

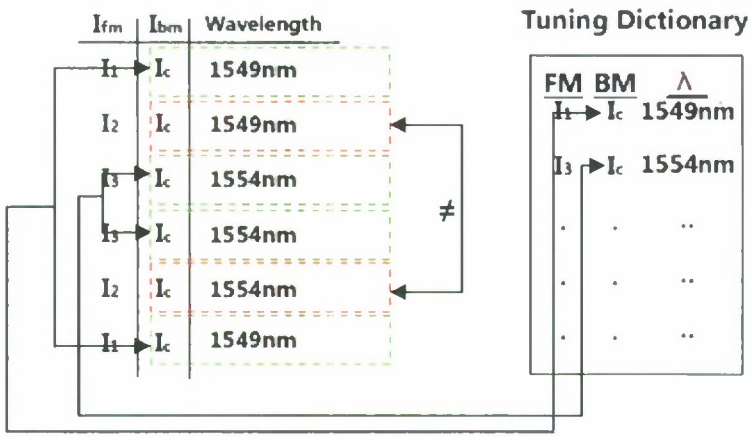


Figure 4.15 - An illustration of the reproducibility algorithm. For a particular back mirror current, if the output for two operating points approached from different directions is the same, the operating point is placed in the tuning dictionary.

However, this leads to some concerns. Does eliminating the points in the hysteretic regions allow enough operating points to perform a tuning ramp? Further more, do these points which are stored in the tuning dictionary allow for a smooth tuning ramp? As explained earlier, the inverse Fourier transform requires constant frequency spacing. Can this spacing be achieved using the points in the tuning dictionary? A tuning ramp was formed using the lookup table method mentioned in Figure 4.15 and the analysis introduced in this section.

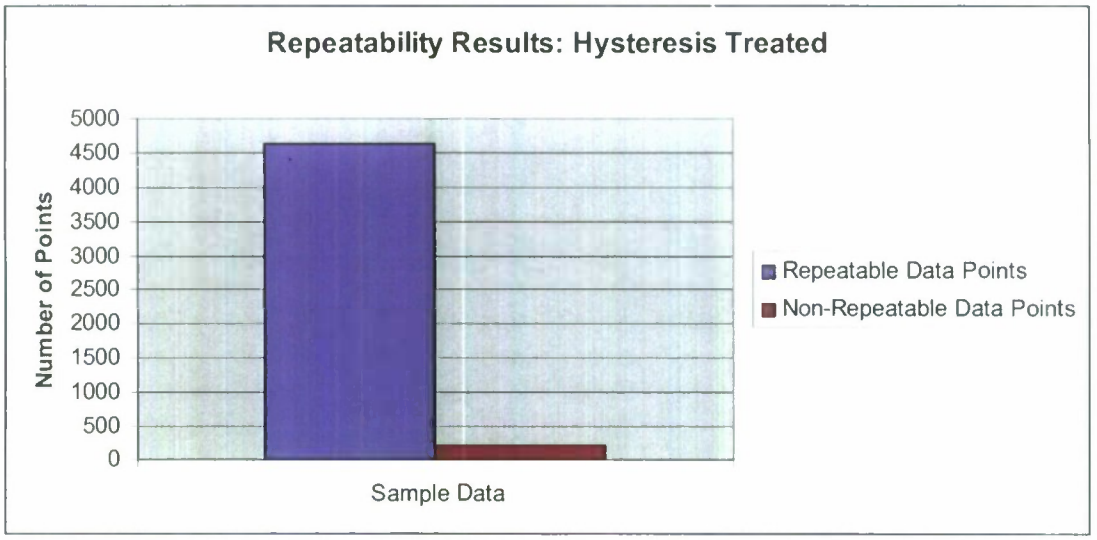


Figure 4.16 - Repeatability results: hysteresis treated. An illustration of the repeatability results after treating hysteresis.

After treating hysteresis, a tuning dictionary was developed. The dictionary was 92% repeatable to .5nm and 81% repeatable to .25nm. The average error can be improved dramatically by paying attention to two areas: phase section tuning and injection current stability.

- Phase section – Tuning the phase section of the laser acts as a fine tuning mechanism. The phase section is capable of tuning the lasers output by .2nm. This allows a method for tuning finer wavelength steps.
- Injection Current Stability - The stability of injection from (the current sources) against temperature was not explored in this report. Using a voltage output D/A, V-I converters were fed voltages from a lookup table. Aspects such as thermal coefficients effect the currents injected into the front and back mirrors. Using programmable current sources which have a more precise output may increase the accuracy of the laser's output wavelength.

In Figure 4.17 the arrows point in the direction where the 5nm jumps occur, which basically point out adjacent modes. Knowing where the mode boundries are in Figure 4.17 yields information about how the modes should be altered by the injection currents in order to produce a linear ramp and avoid hysteresis. After examining Figure 4.17, Figure 4.18 (a) and (b) was developed to show the tuning directions (front and back mirror current values) in which modes jumps occur. Mode jumps were a central topic in preventing repeatability. Therefore it seems that a tuning algorithm would want to avoid crossing the colored boundaries if Figure 4.17 (or do so in a minimal fashion).

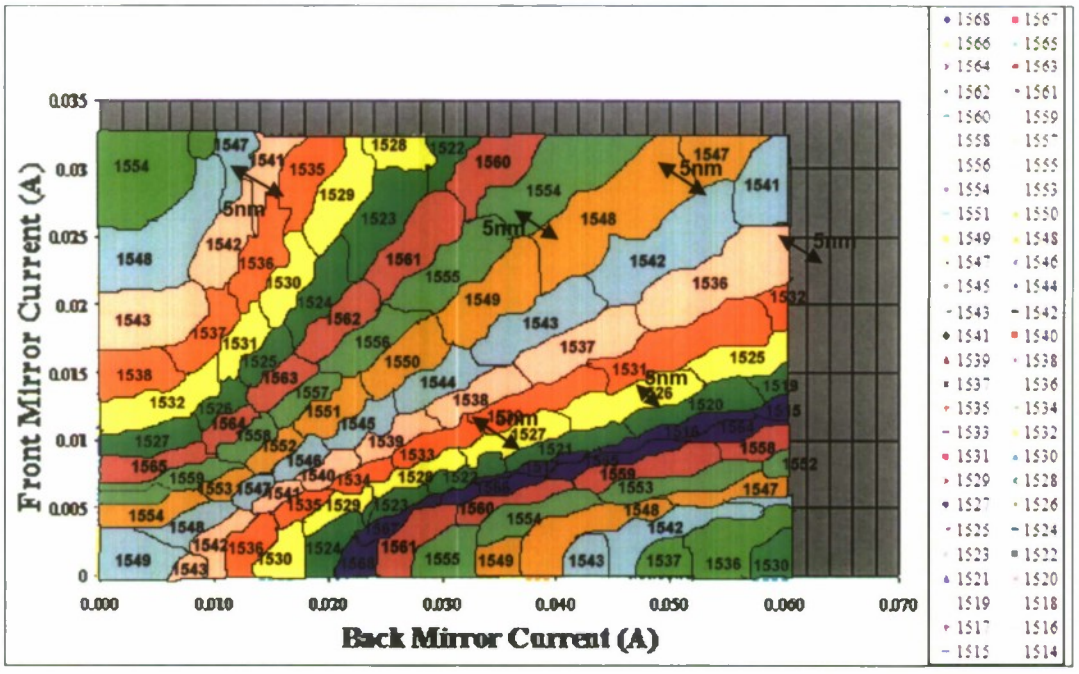


Figure 4.17 - An illustration of where hystesis would oeeur in the tuning eharaeteristies. Since hysteresis only oeeurred during 5nm jumps in wavelength it is assumed that the laser hops on an adjaeent mode with respeet to the mode desired. Eaeh eolor signifies that the seetion of the tuning eharaeteristies lases at a 5nm differenee with respeet to its adjaeent seetions.

It should be noted, that though minimal there are still some non-repeatable points in the final tuning dictionary this may be due to the fact that the hysteresis experiment was not designed to traverse all hysteretic paths. Since it was observed that there was always a 5nm mode jump during hysteresis, the experiment should have been performed in the direction shown in 4.18 (a). The bars in 4.18 (a) are placed there to make the wavelengths readable. All numbers on the graph are in nanometers.

The main problem with the arrows shown in Figure 4.18 (b) (which represent the experiment explained in section 4.3.2) is that they do not traverse all possible hysteretic paths in order to fully treat hystesis. The experiment shown in Figure 4.18 (b) approaches the modal boundaries somewhat obliquely rather than (almost) perpendicularly as shown in Figure 4.18 (a) Hence some hysteresis still occurs causing lasing instability. If one is looking to characterize the hysteresis one should hop modes in the direction shown in Figure 4.18 (a).

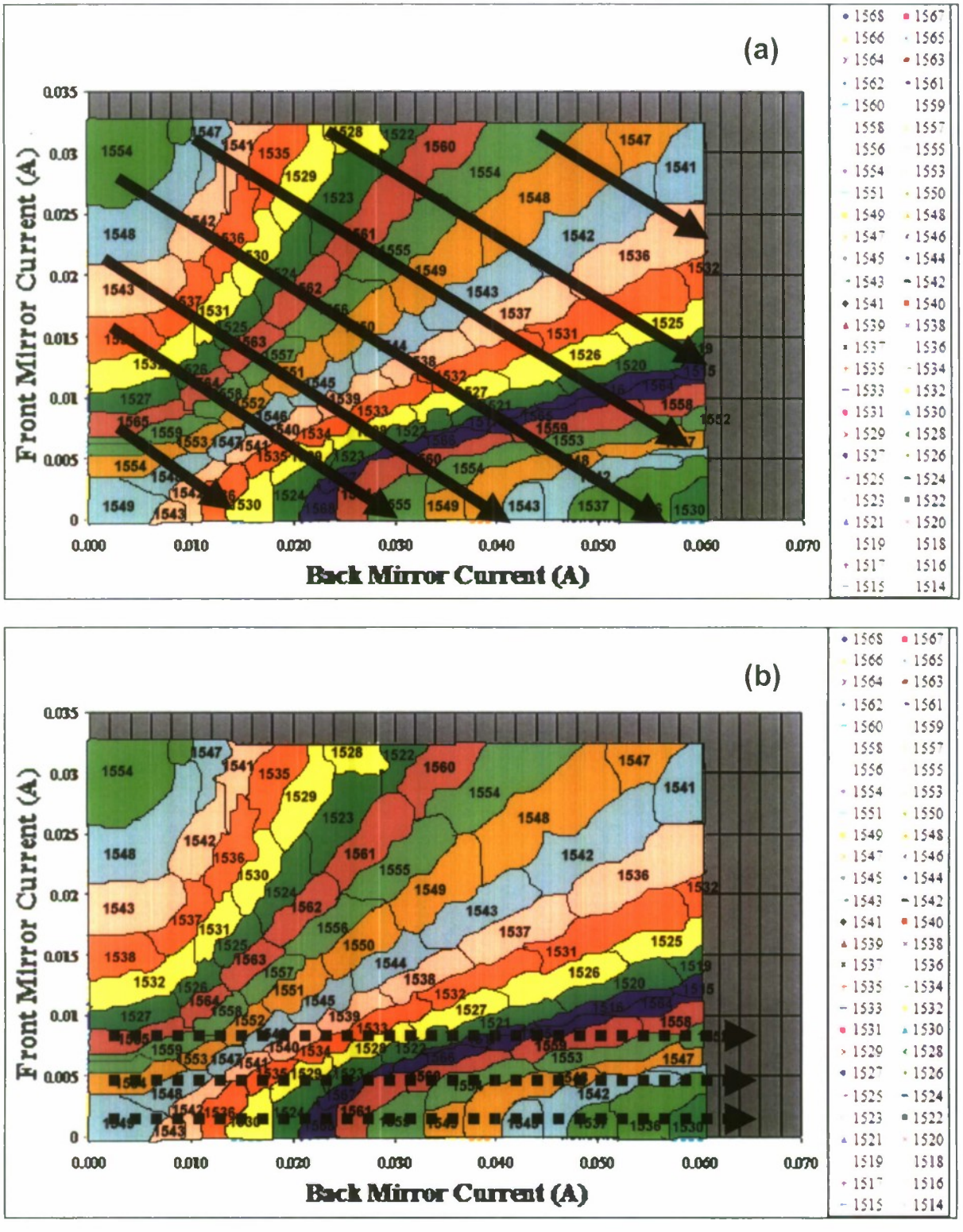


Figure 4.18 - An illustration of the shortcomings of the hysteresis experiment. (a) The tuning path the

hysteresis identification experiment should have taken should have optimized traversal over hysteretic prone (5nm spaced) regions. The bold arrows also are (for the most part) perpendicular to the hysteretic regions. (b) An illustration of the tuning path the hysteresis identification experiment is shown with the dotted arrows. This path does not optimize traversal over hysteretic regions.

Traveling in the direction shown in Figure 4.18 (b) does not fully help solve the repeatability problem during wavelength tuning since the process does not always involves hoping to adjacent modes. Therefore, it makes sense that there was still some unrepeatable points after the analysis discussed in section 4.4.3.

4.5 Switching Speed Accuracy

Both the switching speed and switching accuracy are determined by the surrounding electronics and thermal characteristics[11].

4.5.1 D/A Electronics

The Labjack U12 is a USB DA employed in this application. It only allows for 50 sample/second, therefore keeps the switching time far from the nanosecond regime. The second portion, it the packaged laser used in this applications is intentionally made slower using capacitor. The reason these capacitors are employed is because these lasers are used in communications applications, and one does not want the wavelength to switch due to some system noise.

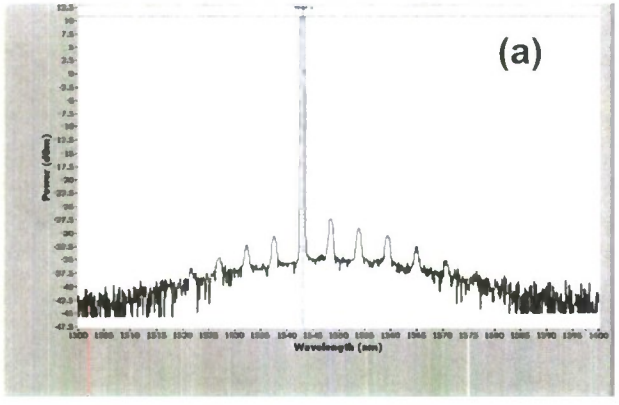
4.6 Thermal Characteristics

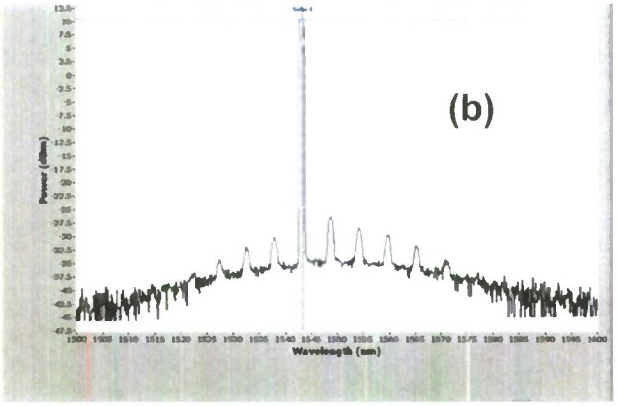
As the temperature of the system shifts, the index of refraction of the mirrors shifts causing a change in output wavelength[11]. This is due to the temperature changes promoting electrons and holes out of the conduction and valance band respectively; changing the electron density.

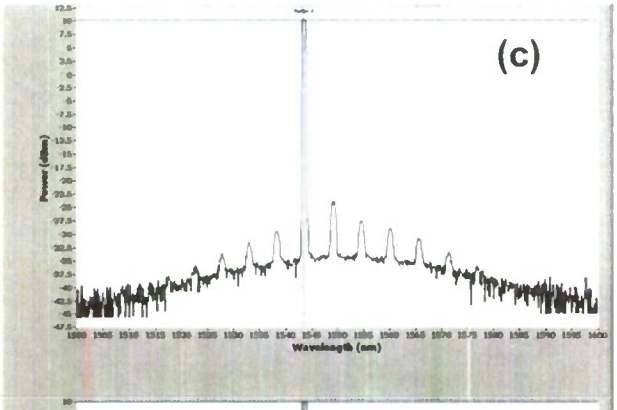
4.6.1 Experiment: Temperature Dependence

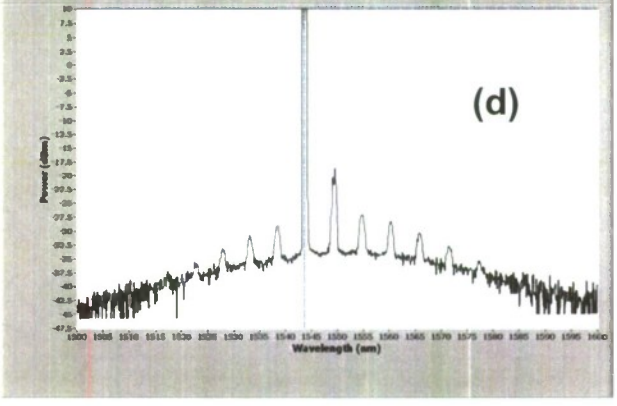
Figure 4.19 shows that a change in current density slides the reflectivity spectrums of the mirrors. Figure 4.19 shows a variation in temperature of the laser while all injection currents into the laser are kept constant. Starting at 16°C (Figure 4.19 (a)) and ramping the temperature of the laser to 24°C, the experiment shows how the output wavelength of the laser is sensitive to changes in temperature. However, Figure 4.19 illustrated the effect of changing the electron density in the conduction band via injection current.

Changing the temperature of the laser should change the reflectivity of each section of the laser simultaneously. This should cause the output to slide. An experiment where, the front mirror, back mirror, gain, phase, and SOA were held to constant values and temperature change was performed. The experiment showed the sliding motion and eventual mode change (shown in Figure 4.19 (c) and Figure 4.19 (d)). This sliding motion signifies a change in carrier density in the mirror, changing the injection current into one of the mirrors also has the same effect on the output of the laser.




(b)
Power (dBm)
Wavelength (nm)


(c)
Power (dBm)
Wavelength (nm)


(d)
Power (dBm)
Wavelength (nm)

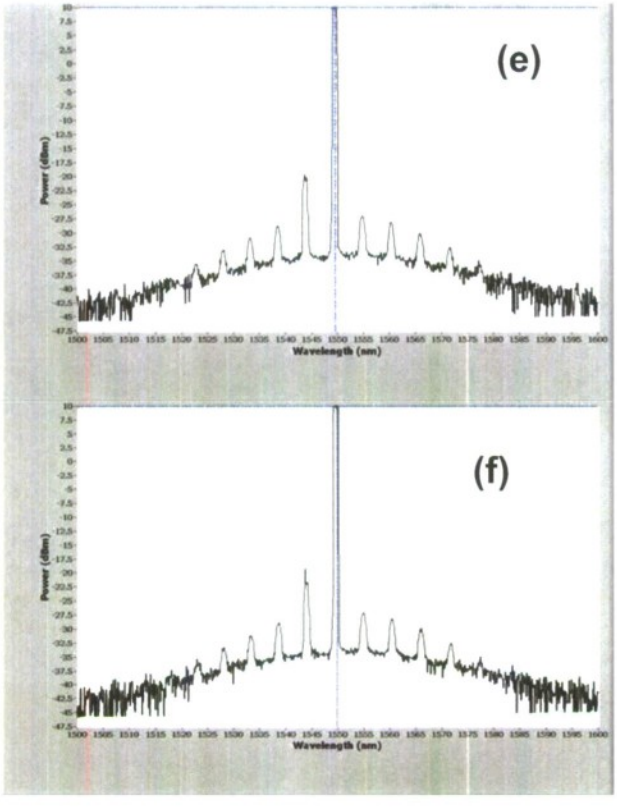


Figure 4.19 - An experiment displaying temperature sensitivity of the Agility 1107. An experiment to illustrate lasing instability results from temperature instability. (a) Agility 1107 spectral output at 16° C. (b) Agility 1107 spectral output at 18° C. (c) Agility 1107 spectral output at 22° C. (d) Agility 1107 spectral output at 23° C. (e) Agility 1107 spectral output at 23.14° C. (f) Agility 1107 spectral output at 24° C.

Notice how the slide, not only slid the lasing frequency of the modes, but also ended up selecting a new lasing mode. This experiment showed the importance to maintaining temperature stability when operating the laser in order to prevent wavelength drift. This experiment also helped conclude that it is impossible to control the laser using an open-loop algorithm without some kind of temperature control.

CHAPTER 5 Strategy for Full Wavelength Tuning

Now that the occurrence of hysteresis has been characterized and the thermal properties of the laser have been noted, an algorithm for implementing a reliable tuning ramp can be hypothesized. This section will discuss the strategies used for open-loop tuning. Specifically, the algorithm for open-loop tuning, protocol to avoid hysteresis, and the method used to determine the lasing wavelength. This section will also discuss repeatability of the tuning ramp

5.1 Strategy for Open-loop Tuning

The strategy for full wavelength tuning involves traversing the laser's characterization curve in a fashion which will the limit large changes in injection current into both mirrors simultaneously, avoid hysteretic regions, and yield a smooth tuning ramp. "Full" tuning implies being able to ramp from the lowest reliable lasing wavelength value to the highest reliable lasing wavelength value.

Figure 5.1 shows the relation between the front mirror and back mirror when both are changing. The positive aspect of Figure 5.1 is that it shows both mirrors don't need to be switched to "random" current values. The case shown in Figure 5.1 is an example where the different colors represent different back mirror voltage values which don't follow a particular pattern while the front mirror maintains a gradual increase in current. The gradual increase in mirror current helps prevent against thermal tuning caused by power dissipation in the mirrors. In communication applications this is protected against by adding capacitors to the laser to block high frequency changes in current. However in OCT applications the use of the capacitor in not desired due to slowdown in tuning operation.

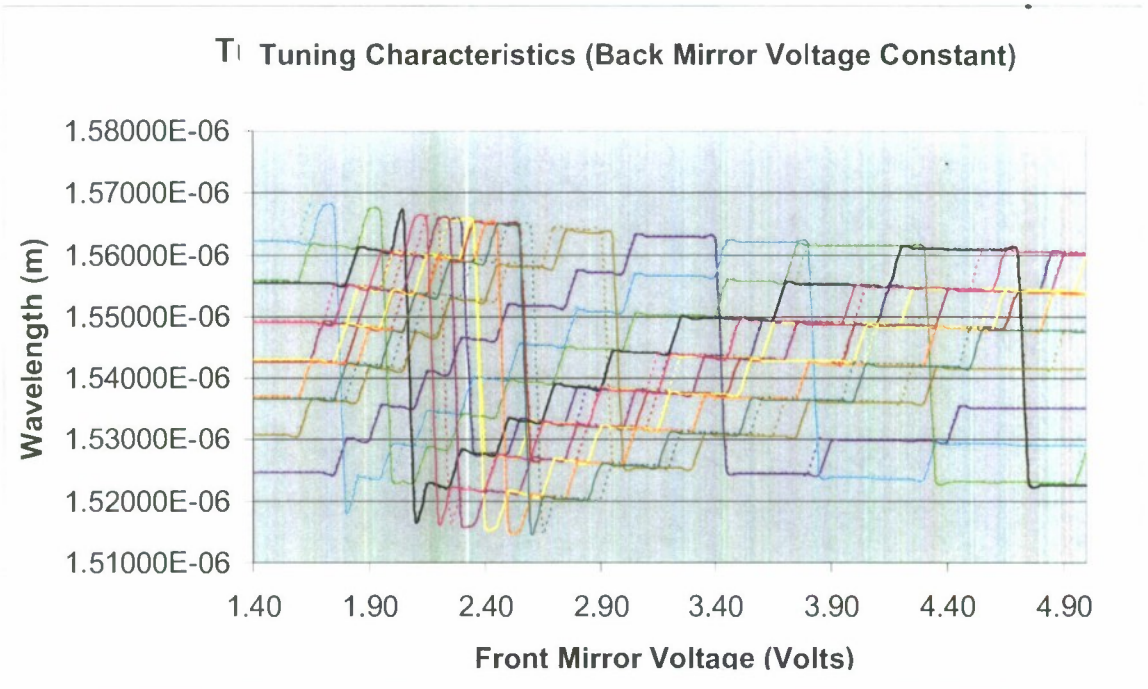
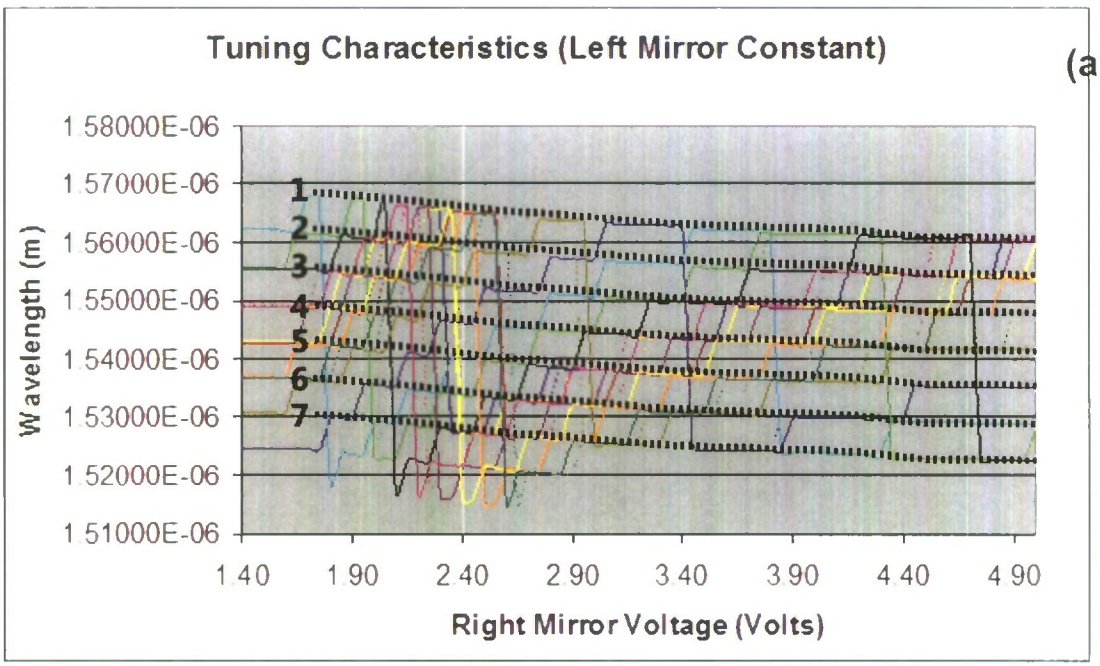


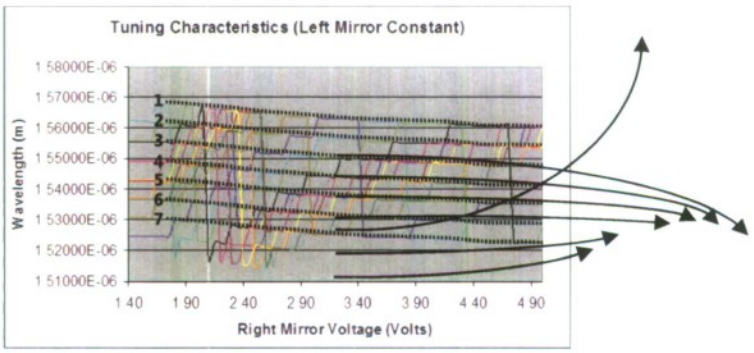
Figure 5.1 - Results from reproducibility experiment. An illustration of the results of the reproducibility experiment defined in section 4.3.2 for various different front mirror values. Plot shows how paths change as both mirrors change.

Figure 5.1 shows different paths the laser takes according to particular constant back mirror settings. Each line represents a particular left mirror voltage while the x-axis represents changes in front mirror voltage. Each color can be seen as a path the laser takes as it ramps through different left mirror and right mirror voltage values.

Figure 5.1 is important in that it displays that a reliable open-loop tuning ramp can be achieved when altered to look like Figure 5.2 (a). Figure 5.2 (a) shows the path needed to be taken in order to ramp the laser's wavelength linearly. This path is shown by the dotted black lines. Figure 5.2 (a) shows the lasing modes intersect at each dotted black line to make up different paths. Therefore, a tuning algorithm will

need to decide which path to jump on in order to produce a tuning ramp.





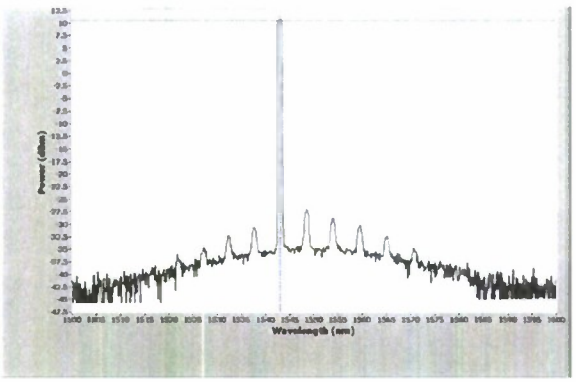
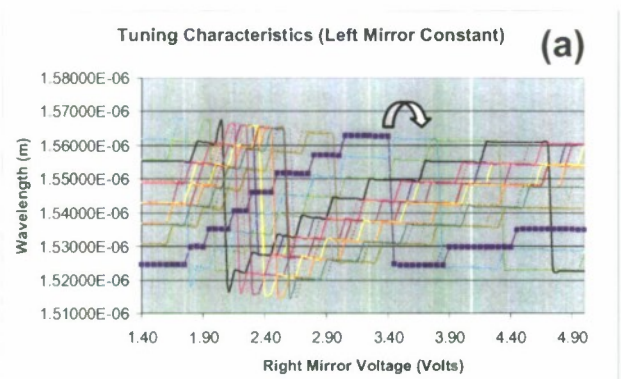
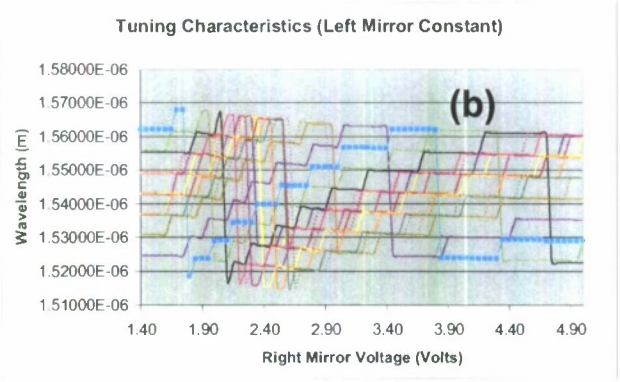


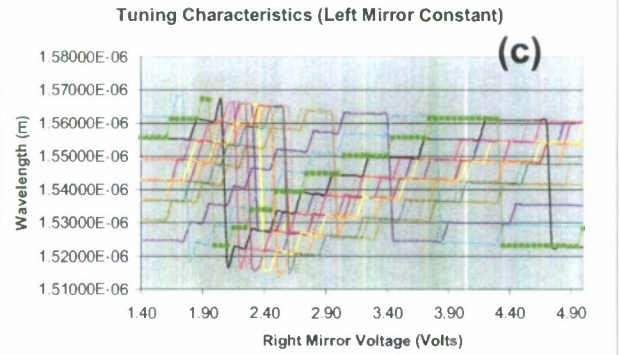
Figure 5.2 - Illustrating where the modes are in the experimental data. (a) Agility 1107 characterization illustrating paths laser travels for different back mirror voltages while front mirror voltage is ramped. (b) An improved Figure 4.4 which takes particular lasing paths into account.

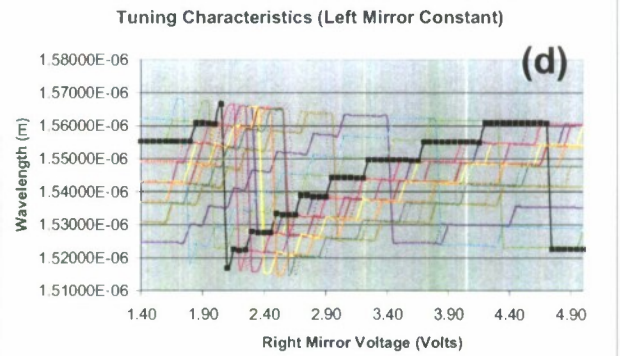
In Figure 5.2 (b) each number next to the black dotted lines (1-7) specifies a particular node. Figure 5.2 (b) can be seen as an improved Figure 4.4 where past and present operating conditions are being taken into account. Like Figure 4.4, Figure 5.2 (b) also shows that the distance between each node is 5nm. This distance represents the spacing between nodes.

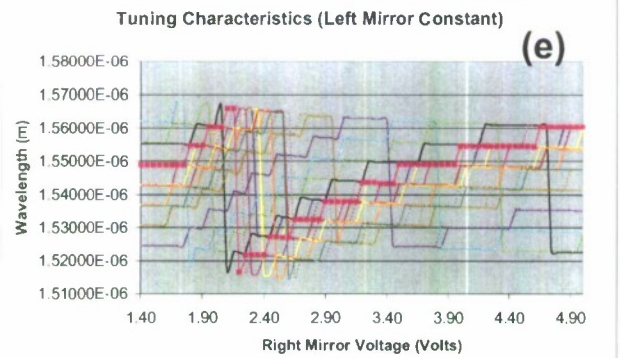
At the hysteretic operating point there is a 5nm difference in lasing wavelength. Notice that each can be tuned to a specific set of wavelength. The algorithm will need to essentially slide the current mode and hop to the adjacent mode.




Tuning Characteristics (Left Mirror Constant)
(b)
Wavelength (m)
1.58000E-06
1.57000E-06
1.56000E-06
1.55000E-06
1.54000E-06
1.53000E-06
1.52000E-06
1.51000E-06
1.40 1.90 2.40 2.90 3.40 3.90 4.40 4.90
Right Mirror Voltage (Volts)


Tuning Characteristics (Left Mirror Constant)
(c)
Wavelength (m)
1.58000E-06
1.57000E-06
1.56000E-06
1.55000E-06
1.54000E-06
1.53000E-06
1.52000E-06
1.51000E-06
1.40 1.90 2.40 2.90 3.40 3.90 4.40 4.90
Right Mirror Voltage (Volts)


Tuning Characteristics (Left Mirror Constant)
(d)
Wavelength (m)
1.58000E-06
1.57000E-06
1.56000E-06
1.55000E-06
1.54000E-06
1.53000E-06
1.52000E-06
1.51000E-06
1.40 1.90 2.40 2.90 3.40 3.90 4.40 4.90
Right Mirror Voltage (Volts)


Tuning Characteristics (Left Mirror Constant)
(e)
Wavelength (m)
1.58000E-06
1.57000E-06
1.56000E-06
1.55000E-06
1.54000E-06
1.53000E-06
1.52000E-06
1.51000E-06
1.40 1.90 2.40 2.90 3.40 3.90 4.40 4.90
Right Mirror Voltage (Volts)

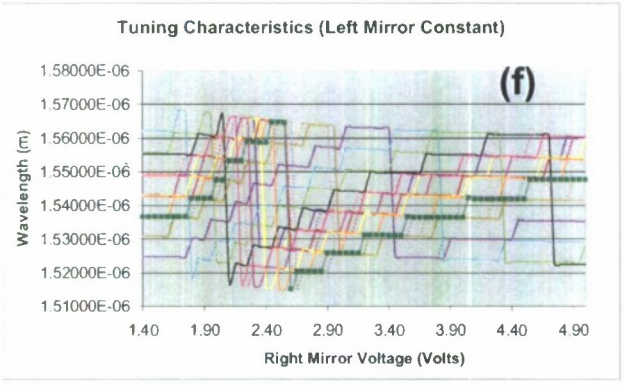


Figure 5.3 - Tuning algorithm. (a) The initial path. (b) The next path the laser must hop to. (c) The next path the laser must hop to. (d) The next path the laser must hop to. (e) Once the mode has reached its limit the algorithm must hop to a different mode. (f) The next path the laser must hop to.

In Figure 5.3 (a) the laser starts on the top most dotted path. It reaches the end of the path and jumps onto the adjacent path pointed to by the arrow in Figure 5.3 (a) which is also the dotted path in Figure 5.3 (b). The laser continues on he dotted path in Figure 5.3 (b) until it hits the end of the path and hops to the path pointed to by the arrow in Figure 5.3 (b) which is also the dotted path in Figure 5.3 (c). The main point of Figure 5.3 (a)-(f) is that the path the laser takes corresponds to a ramp motion with respect to the front mirror.

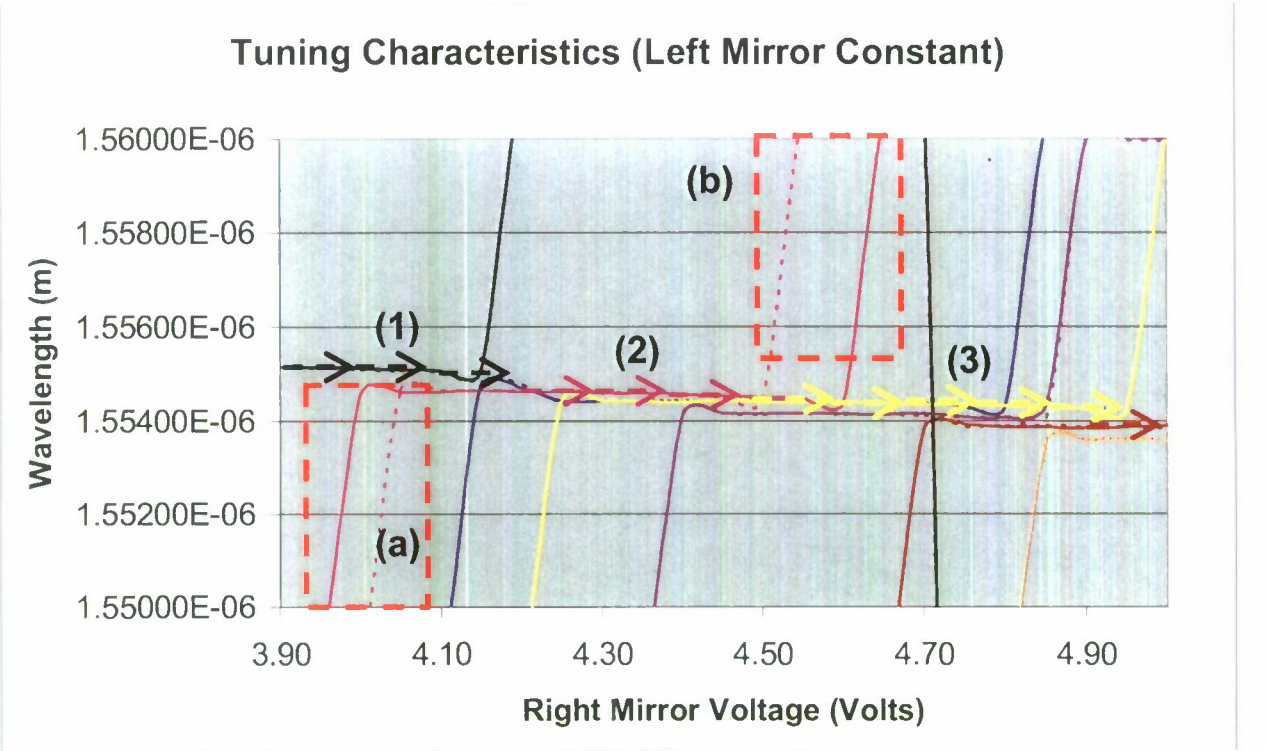


Figure 5.4 - A y-axis close up of the tuning algorithm. A close up of the tuning algorithm showing how the laser paths are chosen to avoid hysteresis. (1) Initial path used to avoid hysteretic region a. (2) Stable path to travel from path 2 to 3. (3) Stable path used to avoid hysteretic region b. (a) Hysteretic region which could affect transfer from path 1 to path 2. (b). Hysteretic region which could affect transfer from path 2 to path 3.

Figure 5.4 shows the factors which decide whether the laser hopes to a new operating path. Jumping to a new operating path involves changing both the front mirror and back mirror current. The dotted squares in Figure 5.4 signify the area where the particular path is unreliable. Therefore, notice that instead of jumping right onto the path surrounded by the dotted square under Path (1) in Figure 5.4, the algorithm continues on Path (1) in Figure 5.4 until a stable portion of Path (2) appears which avoids Hysteretic Region (a) in Figure 5.4. The algorithm will then

switch from Path (1) in Figure 5.4 to Path (2) in Figure 5.4. The algorithm traverses the Path (2) in Figure 5.4 until another unstable region in stumbled upon (Hysteretic Region (b) in Figure 5.4). The algorithm then finds the next reliable path (Path (3) in Figure 5.4) and travels on that path until it turns unstable or the path ends.

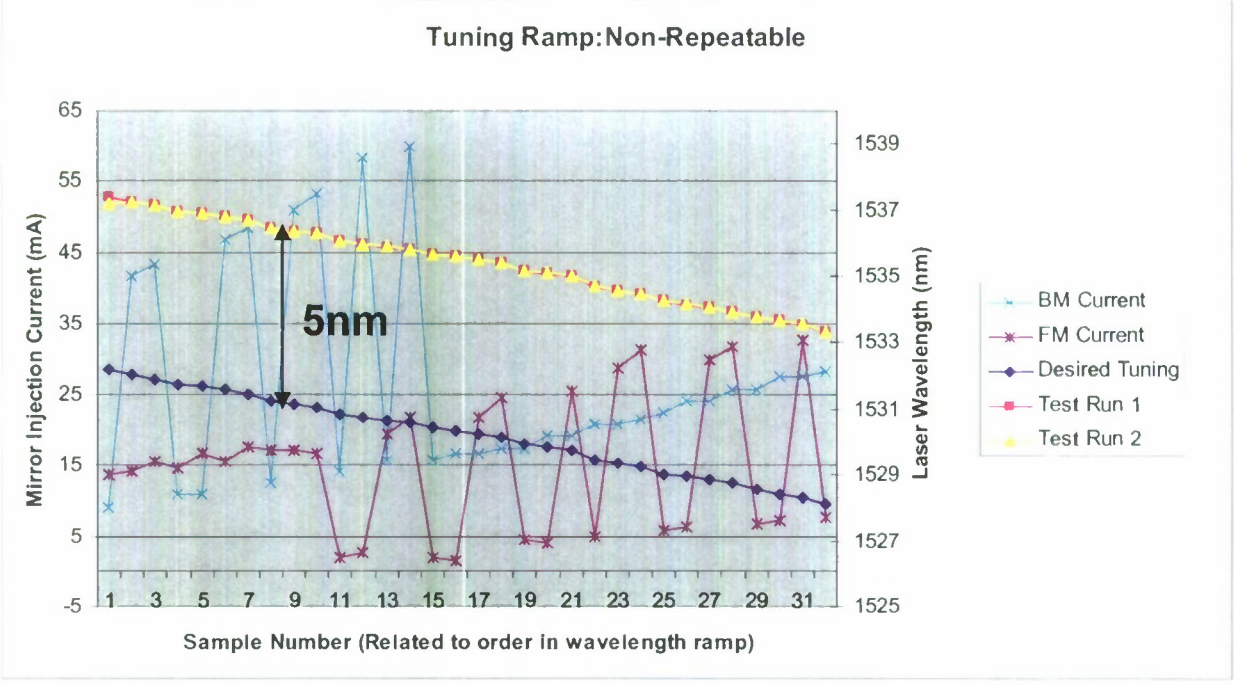


Figure 5.5 - Initial non-repeatable tuning ramp. The figure shows a tuning ramp developed where the back and front mirror currents were not chosen according to the tuning algorithm (which will be presented in Figure 5.7). The tuning ramp exhibited hystesis. "FM Current" and "BM Current" correspond to the left y-axis. "Desired Tuning," "Test Run 1," and "Test Run 2" correspond to the right y-axis.

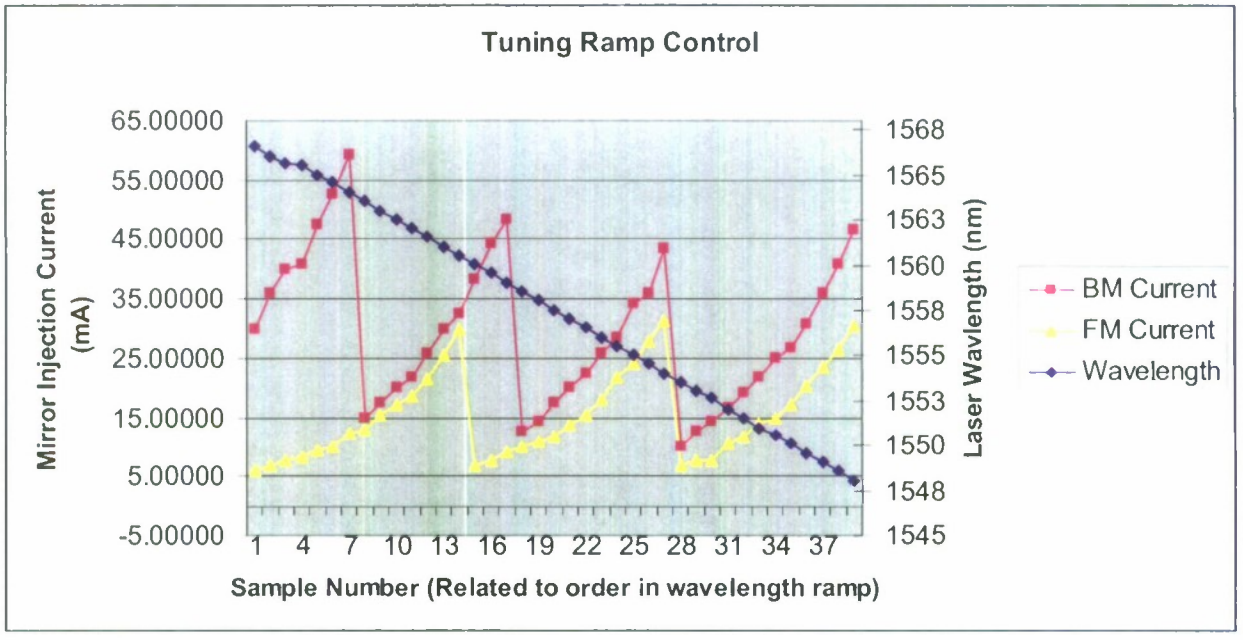


Figure 5.6 - Repeatable tuning ramp. The front mirror and back mirror voltage values were chosen in accordance to the tuning algorithm (which will be presented in Figure 5.7) described in Figure 5.3 and in Figure 5.8. "BM Current" and "FM Current" correspond to the left y-axis. "Wavelength" corresponds to the right y-axis.

Figure 5.6 shows that there is a method to tuning both mirrors in order to produce a tuning ramp. In Figure 5.6 the front mirror voltage increases almost linearly as the back mirror voltage increases in a parabolic pattern. In Figure 5.6 the line marked with triangles represents the lasing wavelength for each set of front and back mirror points. Figure 5.6 differs greatly with the non-repeatable results in Figure 5.5 due to the fact that in Figure 5.5 the injection currents take on more of a less defined pattern.

Figure 5.6 shows that the wavelength changes in a somewhat linear fashion. Note that in Figure 5.6., (for the most part) each mirror's injection current is changing

slowly, leading to few large swings in current density. A large swing in injection current in noted by the discontinuous drops in the graph. Avoiding frequent large swings in injection current helps maintain lasing repeatability. Also, those swings in injection current are a key insight into how the tuning mechanism works and by examining Figure 5.8, one will understand why those swings occur.

Figure 5.6 shows that when the front mirror experiences a large change in injection current, the back mirror increases steadily and vise versa. This gradual change is important to maintain stability and repeatability of the tuning ramp. As a result to Figure 5.6 it shows that there is an ability to write a mathematical algorithm for the tuning ramp. Though this report involves using a fixed lookup table to define a linear ramp, the figure above shows that the same ramp can be constructed mathematically using programmable function generators.

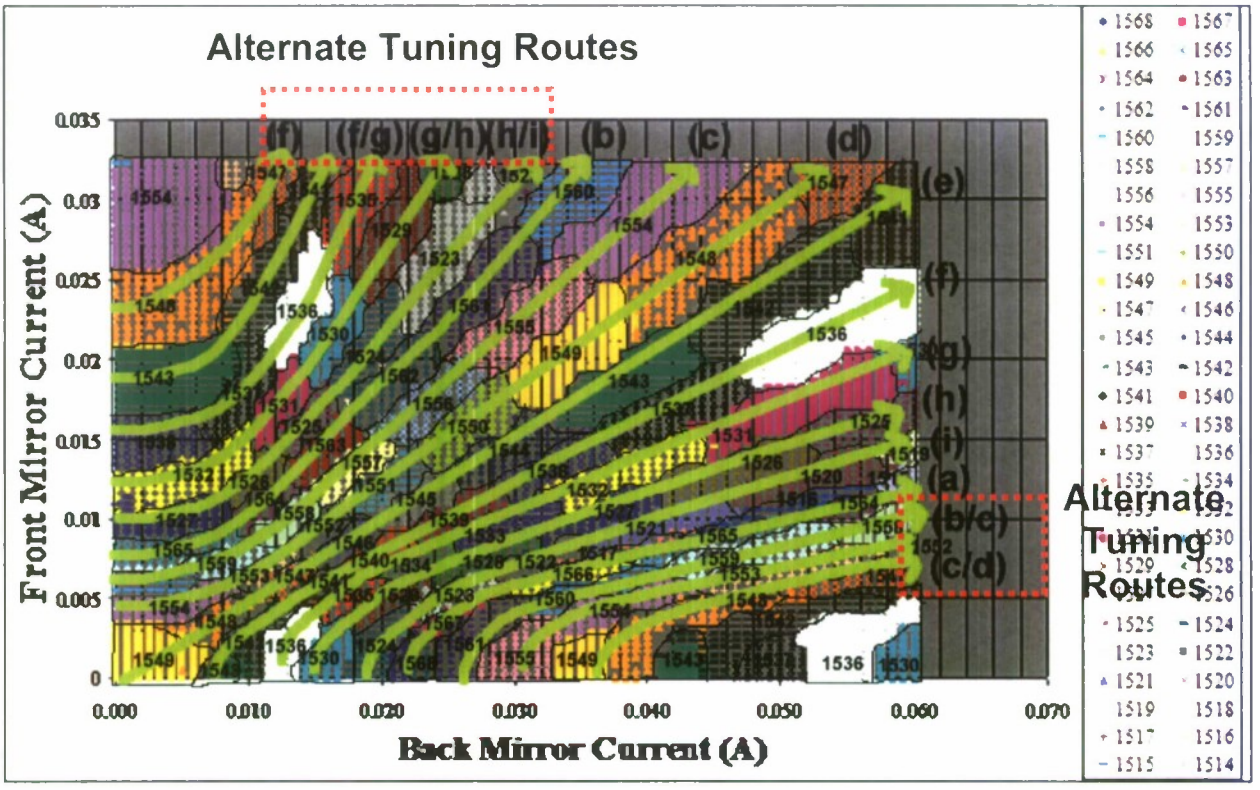


Figure 5.7 - The tuning algorithm. Mirror current control necessary to produce a tuning ramp. (a) Starting path 1568nm – 1564nm. (b) 1564nm – 1560nm. (c) 1559nm – 1554nm. (d) 1552nm – 1547nm. (e) 1546nm – 1541nm. (f)1540nm – 1536nm. (g) 1535nm – 1530nm. (h) 1529nm – 1525nm. (i)1524nm – 1520nm. The letters surrounded by the dotted box represent paths which share the same characteristics as the letters labeling them. These are alternate tuning routes.

Figure 5.6 showed the basic algorithm (or method used to manipulate the front and back mirror injection currents) necessary to produce a continuous downward wavelength ramp from 1566nm to 1520nm. Now that same algorithm will be discussed using the tuning characteristics shown in Figure 5.7. Starting on path (a) in Figure 5.7 one can move in a linear ramp between front mirror current and back mirror current outward. Once reaching the end of the path shown in Figure 5.8 (a) one moves to Figure 5.7 (b) and continues in the same fashion until reaching Figure

5.8 (i). A tuning path in Figure 5.8 can be characterized by the following equation. Basically, increased the front and back mirror currents in a manner which keeps their ratio of change constant will allow for the tuning of a decreasing wavelength ramp. The arrows in Figure 5.7 point in the direction of decreasing wavelength.

Essentially, the discontinuous portions of Figure 5.6 signify tuning path changes in Figure 5.8. For example, switching from path (b) in Figure 5.8 to path (b/c) in Figure 5.8 will cause a discontinuous change in front mirror current, yet during the same path change (path (b) in Figure 5.8 to path (b/c) in Figure 5.8) the back mirror current is ramped linearly.

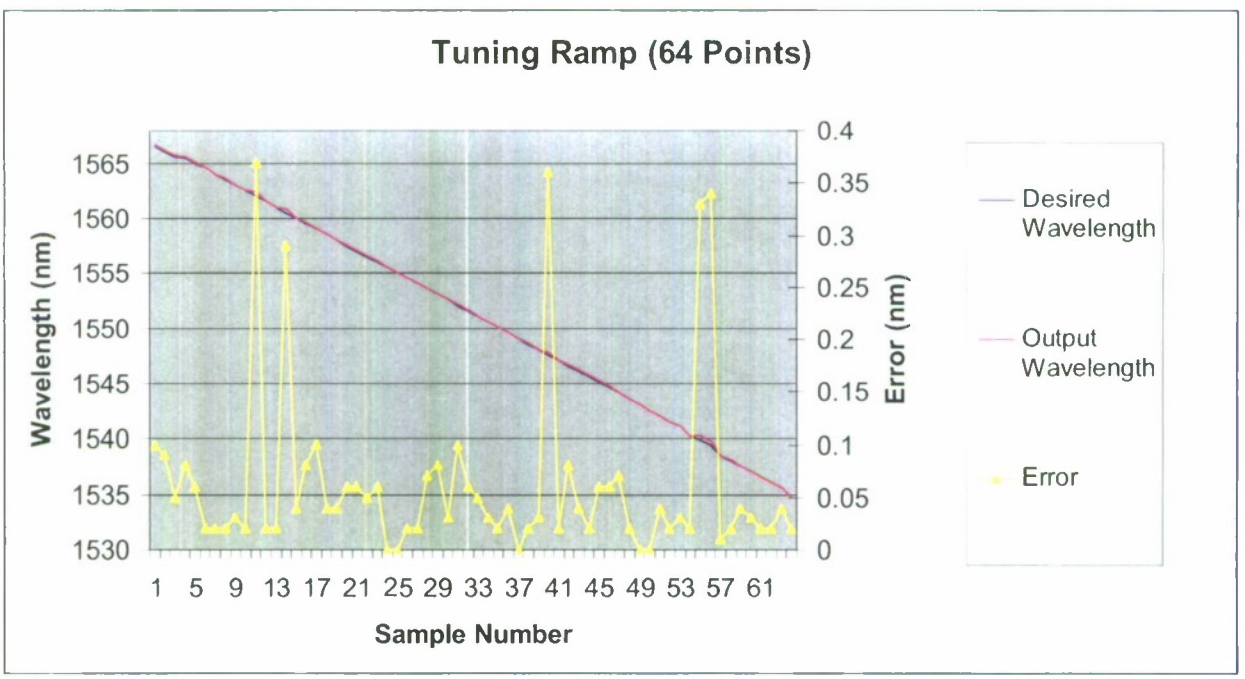


Figure 5.8 - Reproducible tuning ramp results. The desired wavelength ramp was a ramp devolved from data taken 1 week before the output wavelength linear curve. This shows repeatability between the two tests. The tuning ramp above was used in all OCT calculations discussed in the report. "Desired

Wavelength" and "Output wavelength correspond to the left y-axis while "Error" corresponds to the right y-axis.

Figure 5.8 shows data from a repeatability experiment done on a tuning ramp created using the algorithm shown in Figure 5.7. In Figure 5.8, "Desired Wavelength" represents the lasing wavelengths of a test-run upon the initial creation the lookup table used to program the tuning ramp. "Output Wavelength" represents the same lookup table used to create a tuning ramp one week later.

The line marked by triangles in Figure 5.8 signifies the error at each point, the error is signified on the right y-axis. Notice the error is within .5nm. The sources of error for this measurement are due to the fact that there is no feedback mechanism. Also the current sources employed in this application are vulnerable to temperature changes (they will be introduced in Chapter 5) and do not output a consistent current. Using more reliable current sources is a key factor to creating a more reliable tuning ramp.

In conclusion, this section showed that is was possible to create a reliable wavelength tuned ramp using the Agility 1107 Laser.

CHAPTER 6 OCT System Design and Measurements

This chapter will discuss the basics for making OCT measurements and also discuss the system constructed.

6.1 OCT System Design

As stated earlier the OCT system in this report will used the Stepped Ramped Wavelength Architecture. The Stepped Ramped Wavelength Architecture takes advantage of the fast switching capabilities of the SGDBR laser. Using this type of laser as the source component will allow for faster data collection and higher resolution. Figure 6.1 is a schematic showing the chirped laser architecture with an SGDBR laser as the source component. The circuitry for the current control is also included. The numbers in the diagram correspond with the explanations following the figure.

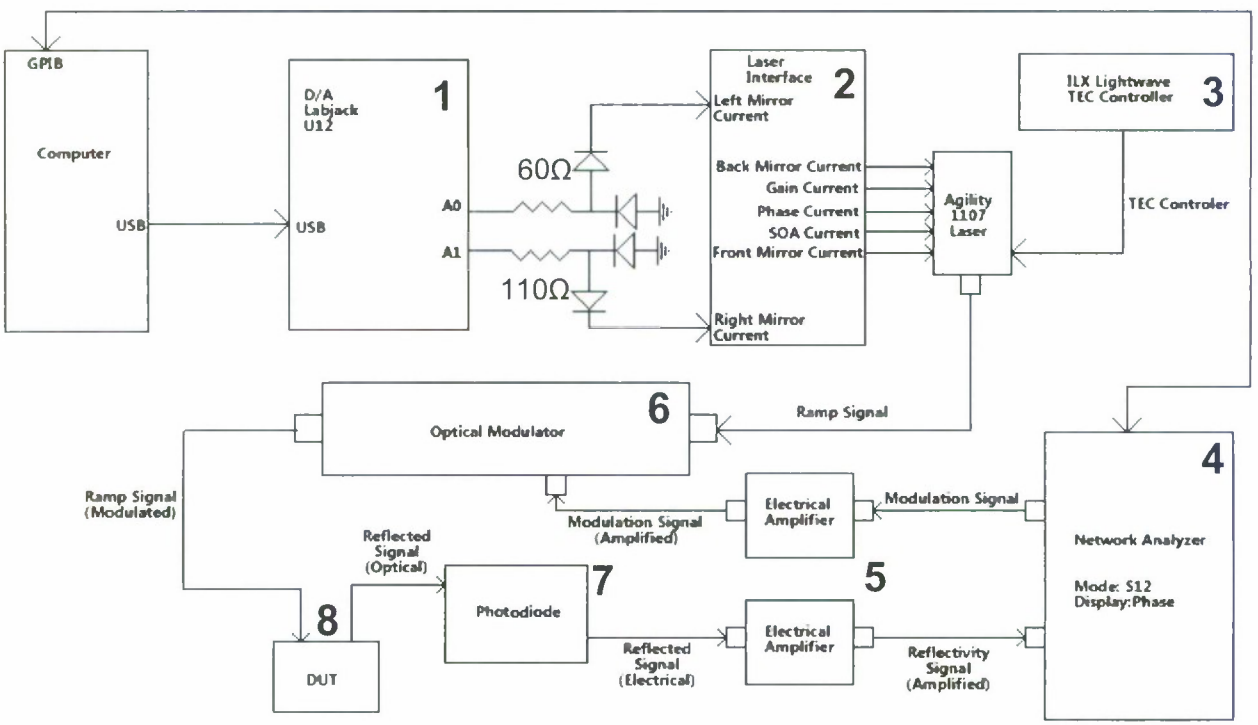


Figure 6.1 - OCT System Schematic.

1. Labjack U12 – The Labjack U12 is a D/A which is used to select injection currents for the front and back mirror of the Agility 1107. The D/A reads voltages from a lookup table stored on the computer designed to output a tuning ramp from the Agility 1107. This is the bottleneck in the system which slows down the injection current tuning leading to the inability of this specific system to take advantage of the fast switching wavelength of the SGDBR laser. In future research it would be more efficient to use a faster D/A or use a current controlled source which can be programmed via GPIB. This would ensure reliable injection currents at a fast tuning rate.

   In between the Labjack are V-I converters which convert the D/A voltage output of the Labjack into currents. Their output can be described by the relation below in (5.1).

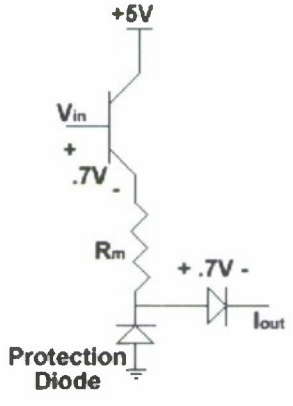


Figure 6.2 - V-I Converter used in this report. (5.1) describes how the circuit converts a voltage to a current.

$$I_{out} = \frac{V_{in} - 1.4}{R_M} \text{A} \tag{5.1}$$

In 5.1 V_{in} is the output value of the Labjack while R_m is the resistor value corresponding to the current conversion for the mirror. The current sources used in this report were not immune to temperature fluctuations. Meaning for a particular input voltage, the output current varied based on temperature. It would be prudent in future research to use a programmable current source as the source components.

Figure 6.3- Labjack U12 D/A. Image from: www.labjack.com

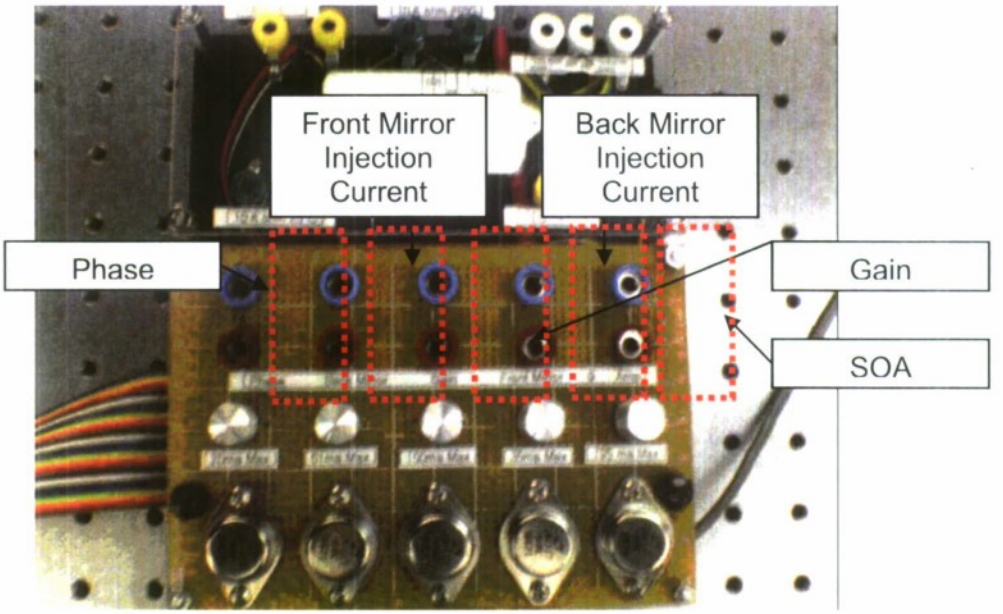


Figure 6.4 – Agility 1107 manual test board. [13]

2. Interface Board – The interface board shown in Figure 6.4 takes currents from the Labjack U12 D/A and provides currents for the phase, gain, and SOA sections of the laser and converts them to a parallel butterfly connection. The problem with this connection is that it is inherently slower than making RF connection strait to the pins of the laser. A fast design would not only make the RF connections strait to the laser's package pins but also by pass the packaged laser all together. Making the RF connections strait to the laser via wire bonds will yield the fastest tuning speed and is an area for further exploration.
3. ILX Lightwave LDC-3700 Series – The ILX Lightwave LDC-3700 provides the TEC control for the laser. This is important for wavelength stability. For all experiments other then temperature dependent experiments, it will be set to 20°C. As shown earlier, the temperature of the laser has a large effect on the

lasing wavelength. The main point to realize is that since the system is an open-loop system, the lookup table for the source needs to be run at the same temperature the lookup table was designed for. Any change in temperature will slide the results or put the laser in hysteresis.

4. Network Analyzer – The network analyzer provides the modulation signal to the optical modulator. It also provides as a tool for phase calculations. The phase and magnitude are then sent to the computer and Inverse Fourier transformed. A computer program then queries the network analyzer for the magnitude and phase of the signal from the DUT. The inverse Fourier transform of the data from the network analyzer can then be taken by using Labview functional elements.
5. Electrical Amplifiers – The amplifiers amplify the electrical signal used to modulate the tuning ramp and also modulate the signal from the DUT. They help provide the proper signal-to-noise ratio so that the measurements can be processed accurately by the network analyzer.
6. Optical Modulator – The optical modulator modulates the ramped signal emitted from the Agility 1107. Figure 6.2 shows a characterization of the modulator with respect to the photo detector voltage. The goal is to operate the optical modulator in its linear region such as to maintain a sinusoidal modulation with the least distortion. Figure 6.2 shows that linear region surrounded by the dotted box in Figure 6.3. The bias used for this report was -1.2 V.

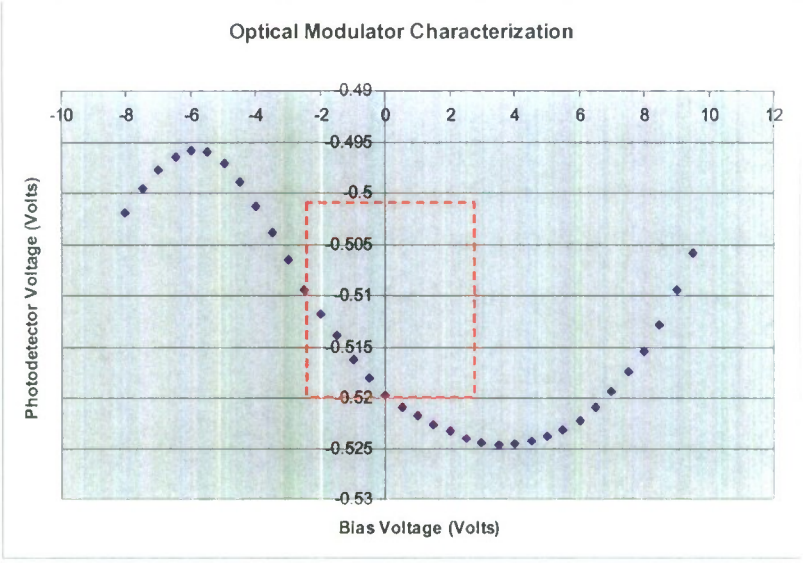


Figure 6.3 - Optical modulator transfer function. The dotted box shows the linear region of optical modulator.

7. Photodiode – The photodiode converts the reflected optical signal from the DUT to an electrical signal so that it may be analyzed by the electrical network analyzer. It was found that if the photodiode was biased, it helped increase the output current induced by the photons from the DUT. The reverse bias value used for this report was 5 volts.
8. DUT – Device Under Test. Two devices were tested to display functionality of the OCT system. The first device was a fiber optic cable; the second device was a fiber optic interferometer. The results of the two tests will be discussed later.

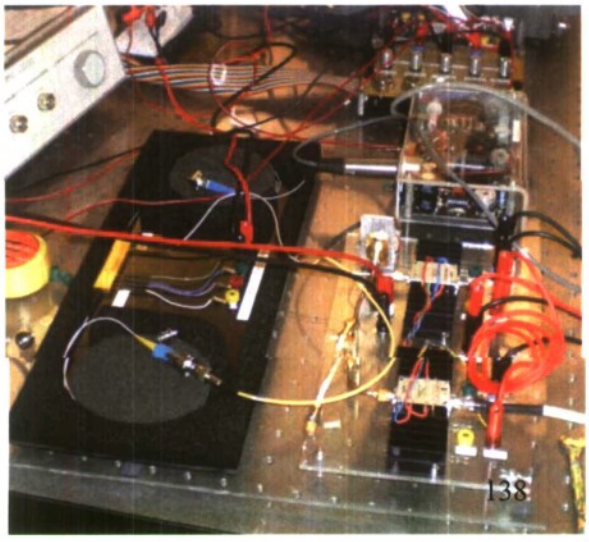


Figure 6.4 - A pieture of the modulation/amplifier subsystem.

6. 2 Basic Measurement

There are two steps in the signal processing the first is producing the modulated signal, the second is processing the magnitude and phase from DUT. The first step involves the taking data in the Fourier Domain.

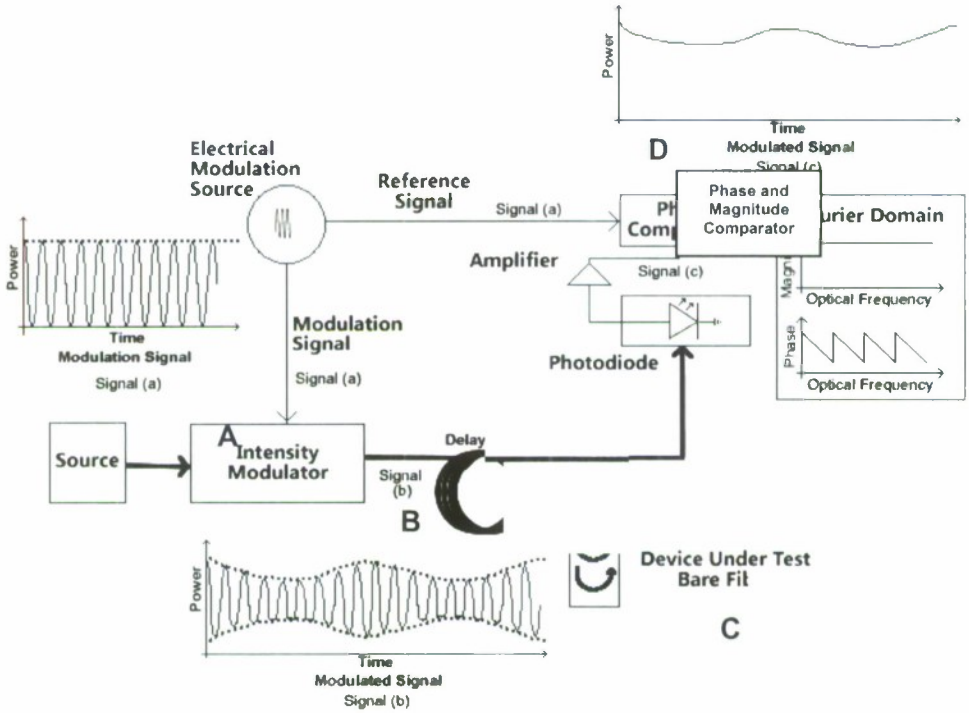


Figure 6.5 - The setup for a basie OCT measurement. (Figure 2.9 shown again).

A. First a wavelength is emitted from the laser and is modulated by a constant modulation frequency (130MHz was used in this report). The equation for this calculation is shown in (2.2).

B. The optical signal is modulated by the optical modulator. This modulation allows a reference to compare points in the Fourier domain. The comparison

is done internally within the network analyzer. The magnitude and phase are queried from the network analyzer at the end of the measurement.

C. The signal strikes the DUT. The DUT can range to anything from a fiber, to body tissue, to mechanical components. Anything which reflects/transmits within the lasing region of the laser allows for the processing of data and an output OCT image.

D. The signal from the DUT is processed by the network analyzer. The network analyzer then the return a magnitude and phase to be processed by a computer program.

6. 3 Signal Processing

The signal processing of the data was done using Labview. A hamming window was used to filter the data and the inverse Fourier Transform was taken in order to develop a time domain response. The program employing this can be seen in Appendix D.

6.4 Measurement Data

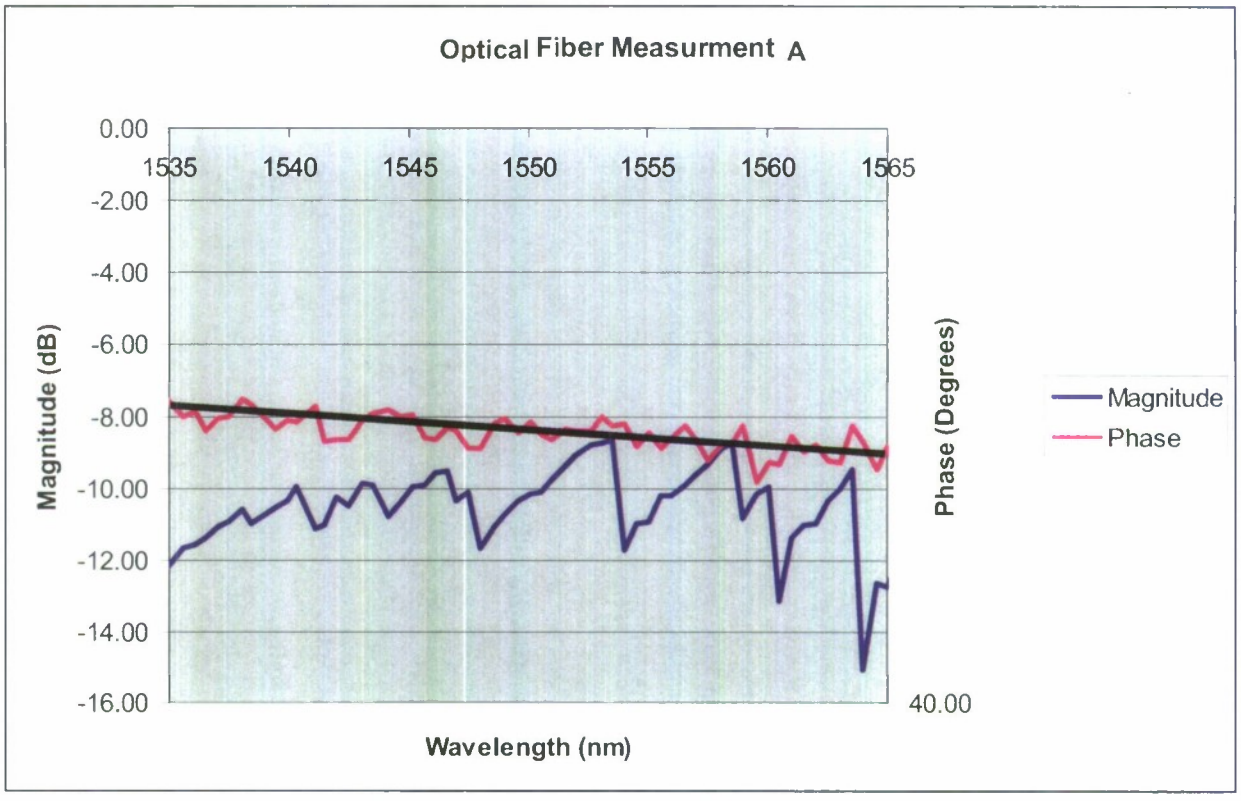


Figure 6.6 - Optical Fiber Measurement: Test Run A.

Notice that the magnitude is for the most part constant while the phase is experiencing a somewhat linear change in slightly phase downward. The constant magnitude and linear change in phase is expected for a fiber. The longer the fiber the steeper the slope of the phase change. The slope of this phase change corresponds to the group delay of the system. One important aspect to note is that since the OCT system measures a change in group delay, a very long fiber is necessary to measure a significant change in group delay. This change in phase becomes more prevalent in Figure 6.6.

Notice in Figure 6.6 that by zooming into the phase section a linear change in phase

can be seen. The goal of this experiment is to show proof of functionality of the OCT system. Some changes can be made to improve resolution such as varying the wavelength step size. For longer fibers a smaller step size is needed since the change in phase is more noticeable since the system has a greater group delay. Since the phase change was so small in Figure 6.6, a longer fiber was used to obtain the data shown in Figure 6.7.

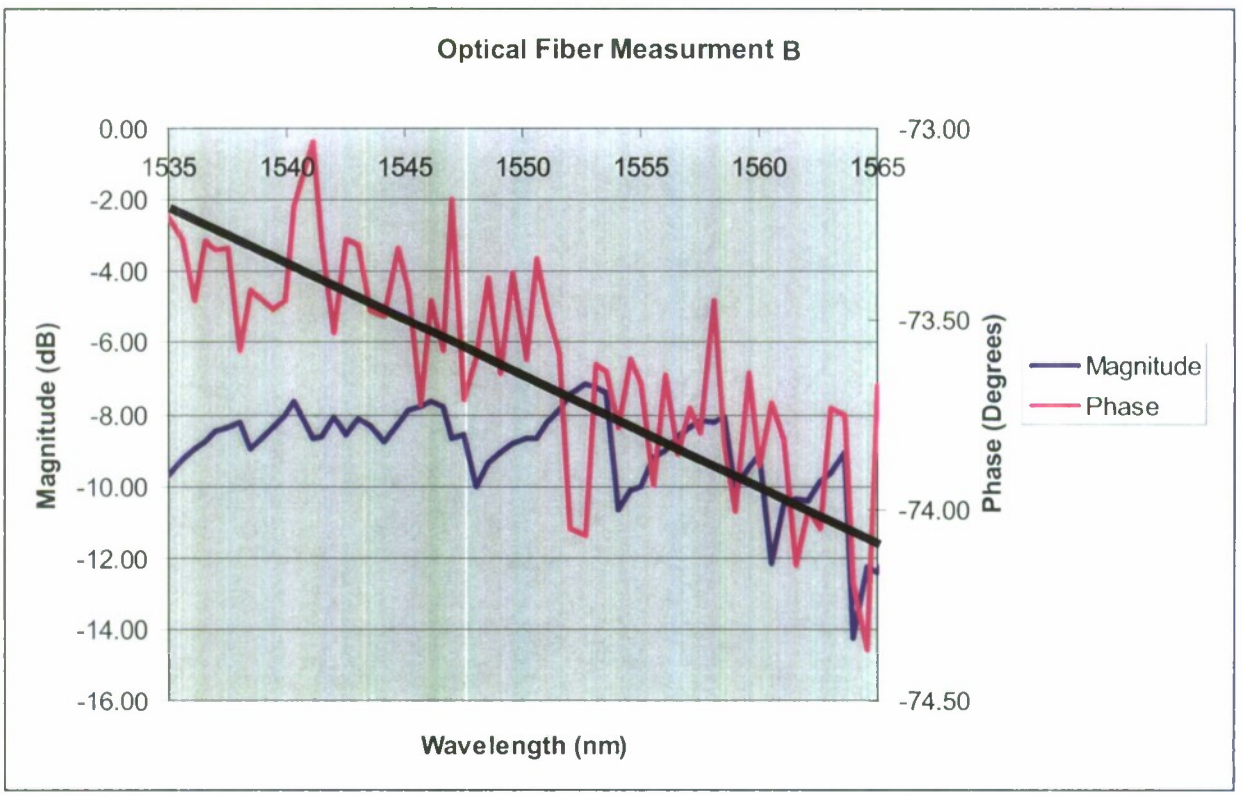


Figure 6.7 - Optical Fiber Measurement: Test Run B.

The small changes in phase makes sense since the system measures a change in group delay and that change is very small when one changes the frequency of the envelope. Notice the black line shows that the phase changes in a linear fashion. This is expected in a fiber. The reason the experiment could not further demonstrate

the measurement of the length of a fiber was that the optical modulator used in this experiment did not produce a modulation with a signal-to-noise ratio high enough.

CHAPTER 7 Conclusion and Future Work

This report showed that it is possible to develop an open-loop tuning ramp using the Agility 1107 CW Widely Tunable Laser. This open-loop tuning research provides the framework for further research in OCT in real-time high resolution OCT applications.

It was shown in Chapter 5 that the Agility 1107 laser can be tuned in a linear ramp fashion from 1520nm to 1565nm. The 6THz frequency resolution obtained from these results is acceptable in OCT applications. In showing that the laser can be tuned over 1520nm - 1565nm a tuning algorithm was presented in order to create a repeatable linear tuning ramp using the Agility 1107 laser. This algorithm involved exploring and mitigating the factors which effected wavelength repeatability of the laser (e.g. hysteresis). These developments led to a tuning algorithm which involved minimal mode changes, and took advantage of the mode sliding capabilities of the laser.

The analysis presented in this report showed that SG-DBR lasers are promising in OCT. SG-DBR lasers have the potential of creating smaller OCT real time systems due to the laser's small size and fast switching capability . The work done on SGDBR laser tuning helps contribute to the understanding of the wider application of these lasers outside the field of communications which does not allow the laser to be used to its fullest extent. It was shown that the lasers can be tuned in an open-loop fashion and produce a tuning ramp.

The main work of this report showed that tunability of SG-DBR Laser is rather

complicated however there is a method to obtaining a tunable ramp. Figure 5.8 showed that tuning two mirrors in conjunction with one another can allow for a consistent mathematically programmed method for wavelength tuning.

Future work can go into making the system both faster and more reliable with respect to wavelength tuning. Understanding the different portions of the laser is important to developing accurate tuning. Specifically, characterizing how the phase section effects the overall tuning algorithm.

The main work that should be done in the future is raising the speed of the system. For example using an unpackaged version of the laser or using a packaged version which does not have the capacitance disadvantages of the Agility 1107 will allow for faster wavelength switching. Also, making RF connections directly to the laser is essential in yielding the fastest laser switching output. It should be noted that also in order to take advantage of the compact size and the low power consumption traits of the laser, an unpackaged laser implementation is ideal due to its speed lack of hindering capacitors.

Work can also be done on creating an improved compact thermally independent current source. Since the overall goal is to develop a compact design, creating a circuit board where the current controllers output a reliable current regardless of temperature value will be important.

In closing, this report presented an argument that SG-DBR lasers can be used in

OCT applications. During that argument, the discussion showed that SG-DBR lasers are not only capable of being used in OCT, but also seem to be a very promising for advancing current OCT technology. With various advancements to the system presented in this report, a high speed real-time OCT system can be developed using SG-DBR lasers.

Graduate Students: Andrew Dekelaita (MSEE Dec. 2007), Mike Bernacil (Anticipated MSEE June 2008), Shane O'Connor (Anticipated MSEE June 2008), and Ben Maher (Anticipated MSEE, June 2008).

Senior Project Students: Josh Anthony (MEMs Beam Steering Apparatus – graduated June 2007), Ting Kung, (Nortel Tunable Laser evaluation- Graduated June 2007)), Sam McGinn (optical modulator assembly- Graduated June 2007), and Joe Cassio (voltage to current converter circuitry Graduated June 2007).

Bibliography

[1] K. V. Larin, M. S. Eledrisi, M. Motamedi, R. O. Esenaliev, *Noninvasive Blood Glucose Monitoring with Optical Coherence Tomography,* vol. 25, Diabetes Care, 2002, p. 2263

[2] Thorlabs, "Optical Coherence Tomography" (White Paper, Thorlabs)

[3] Tycho, Andreas, "*Optical Coherence Tomography: Monte Carlo Simulation and Improvement by Optical Amplification*" (Ph.D. Dissertation, Technical University of Denmark, 2002) p.1, 2, 4

[4] D. Derickson, "White Paper: Optical Vector Network Analysis Applications Enabled by Fast-Tuning Single-Chip Wavelength Tunable Lasers" (White Paper, California Polytechnic University)

[5] H. G. Bukkems "New Approaches to Widely Tunable Semiconductor Lasers" (Ph.D. Dissertation, Technische Universiteit Eindhoven, 2006) p. 27, 36

[6] J. E. Simsarian, M. C. Larson, H. E. Garrett, H. Xu, T. A. Strand, *Less Than 5-ns Wavelength Switching With an SG-DBR Laser*, vol. 18, IEEE Photonics Technology Letters, 2006, p. 565

[7] J. Buus, E. J. Murphy, *Tunable Lasers in Optical Networks*, vol. 24, Journal of Lightwave Technology, 2006, p. 5

[8] L. A. Coldren, G. A. Fish, Y. Akulova, J. S. Barton, L. Johansson, C. W. Coldren, *Tunable Semiconductor Lasers: A Tutorial,* vol. 22, Journal of Lightwave Technology, 2004, p. 193

[9] G. Sarlet, G. Morthier, R. Baets, *Control of Widely Tunable SSG-DBR Lasers for Dense Wavelength Division Multiplexing*, Journal of Lightwave Technology, 2000, p.1128

[10] M. C. Larson, M. Bai, D. Bingo, N. Ramdas, S. Penniman, G. A. Fish, L. A. Coldren, "Mode Control of Widely Tunable-Laser SG-DBR Lasers" (White Paper: Agility Communications)
http://www.ece.ucsb.edu/Faculty/Coldren/reprints/2002Reprint/papers/larson_mode.pdf

[11] J. E. Simsarian, A. Bhardwaj, J. Gripp, K. Sherman, Y. Su, C. Webb, L. Zhang, M. Zirngibl, *Fast Switching Characteristics of a Widely Tunable Laser Transmitter*, vol. 15, IEEE Photonics Technology Letters, 2003, p.1038

[12] S. McGinn, "Modulator and Amplifer Subsystem for a Tunable-Laser Based Optical Tomography Measurement System" (Senior Project, California Polytechnic University, San Luis Obispo, 2007)

[13] M. Digino (Senior Project, California Polytechnic University, San Luis Obispo, 2007)

[14] B. Choi, T. E. Miller, J. Kim, J. N. Goodman, G. Vargas, G. Aguilar, J. S. Nelson, *Use of Optical Coherence Tomography to Monitor Biological Tissue Freezing During Cryosurgery*, vol. 9, Journal of Biomedical Optics, 2004, p.282

[15] J. R. Taylor, C. D. Zafirator, M. A. Dubson, Modern Physics for Scientist and Engineers. New Jersey: Prentice Hall. 2004. p. 353

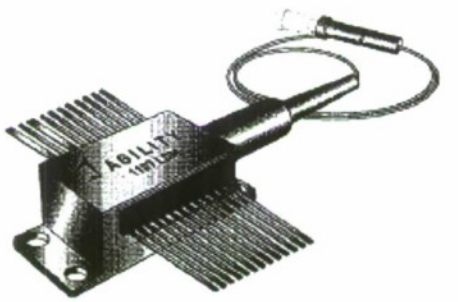
[16] L. Yu, Z. Chen, *Improved topographic imaging of wavelength scanning digital holographic microscopy by use of digital spectral shaping*, vol. 15, Optics Express, 2007, p. 878

[17] T. L. Troy, S. N. Thennadil, *Optical properties of human skin near infrared wavelength rand of 1000 to 2000 nm*, Journal of Biomedical Optics, 2001, p.167

Appendix A: Agility 1107 Data Sheet

AGILITY

Preliminary Product Data Sheet
March 14, 2003



Agility 1107 CW Widely Tunable Laser Diode Module

Description

The Agility 1107 Widely Tunable Laser Diode Module is designed for use in applications where direct access to the tunable laser chip is required. This module supports a number of diverse design requirements for telecommunications and scientific applications. The compact footprint allows for small, flexible designs and a high degree of customization.

Applications

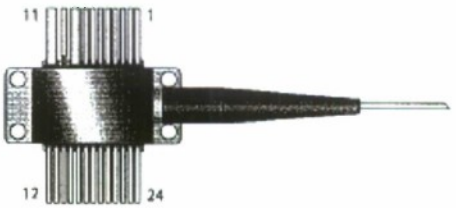
- Optical Networks
- Test and Measurement
- Sensors
- Avionics
- Fast Wavelength Switching

Features

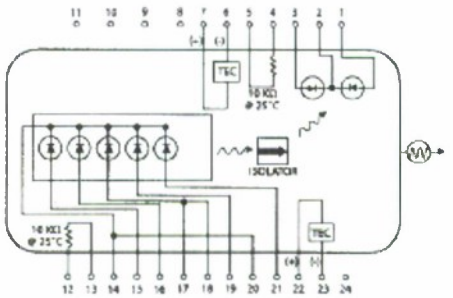
- Sampled-grating distributed Bragg reflector (SG-DBR) widely tunable laser chip
- Hermetically-sealed 24-pin butterfly package with optical isolator and wavelength locker
- Capable of covering extended C-band (1528-1568 nm)
- Minimum output power of 10 mW over entire C-band
- Monolithically integrated semiconductor optical amplifier (SOA)
- Side mode suppression ratio > 37 dB
- Relative intensity noise < −135 dB/Hz
- Linewidth < 10 MHz
- Wavelength locker includes etalon and reference photodiodes
- Integrated TECs and precision thermistors
- Single-mode polarization maintaining fiber pigtail

Pin Descriptions

Pin	Description
1	Etalon Photodiode Anode (–)
2	Photodiode Common Cathode (+)
3	Reference Photodiode Anode (–)
4	Locker Thermistor
5	Locker Thermistor
6	Locker TEC (–)[a]
7	Locker TEC (+)[a]
8	NC
9	NC
10	NC
11	NC
12	Laser Thermistor
13	Laser Thermistor
14	Laser Cathode (–)
15	Back Mirror Anode (+)
16	Phase Anode (+)
17	Gain Anode (+)
18	Gain Anode (+)
19	Front Mirror Anode (+)
20	Laser Cathode (–)
21	SOA Anode (+)
22	Laser TEC (+)[a]
23	Laser TEC (–)[a]
24	NC



Pin Numbers



Schematic of the Laser Diode Module

a. Positive current through the laser or locker TEC cools the laser or locker.

Package Absolute Maximum Ratings

Parameter	Symbol	Min	Max	Unit
Package Case Operating Temperature	T_p	0	+70	°C
Storage Temperature	T_s	–40	+85	°C
Lead Soldering Time at 250°C			10	sec
Package Mounting Torque	T		20	In-oz

Optical and Electrical Characteristics

Parameter	Symbol	Min	Typ	Max	Unit
Laser Threshold Current [a]	I_{TH}		25	40	mA
Optical Power Out at Any Wavelength [b]	P	10			dBm
Tuning Range	λ	1528.384 191.10		1568.773 196.15	nm THz
Operating Currents					
• SOA	I_{SOA}			195	mA
• Front Mirror	I_{FM}			35	mA
• Gain	I_{GN}			150	mA
• Phase	I_{PH}			20	mA
• Back Mirror	I_{BM}			60	mA
Tuning Efficiency (Active)					
• SOA			0.03		GHz/mA
• Gain			0.3		GHz/mA
Tuning Efficiency (Passive)					
• Front Mirror			40		GHz/mA[d]
• Phase			35		GHz/mA[d]
• Back Mirror			30		GHz/mA[d]
Linewidth (3 dB) [c, d]	LW			10	MHz
Relative Intensity Noise [c, e]	RIN			-135	dB/Hz
Side Mode Suppression Ratio [c, f]	SMSR	37			dB

Note: The butterfly-packaged laser will support a minimum output power of 10 mW over the 1528.384–1568.773 nm range. Wavelength and power calibration are the responsibility of the user. The control circuits used will have a significant impact on laser performance and reliability.

a. Measured at T_{LASER} = 22°C, I_{PH} = 0 mA, I_{FM} = 0 mA, I_{BM} = 0 mA, I_{SOA} = 150 mA
b. Over lifetime and case temperature range of 0–70°C
c. Measured at mode center with T_{LASER} = 22°C, I_{PH} = 4 mA, I_{GN} = 150 mA, I_{SOA} = 150 mA, 1 mA < I_{FM} < 35 mA, 2 mA < I_{BM} < 60 mA
d. Effective 3 dB linewidth corresponding to white phase noise density measured using coherent discriminator method
e. Measured at worst-case between 0.1 and 10 GHz
f. 0.1 nm resolution bandwidth

Wavelength Locker Characteristics

Parameter	Symbol	Min	Typ	Max	Unit
Locker Etalon Free Spectral Range	FSR	49.97		50.03	GHz
Etalon Temperature Sensitivity			1.25		GHz/°C
Locker Temperature Setpoint	T_{LOCK}	44		70	°C
Reference Detector Photocurrent to Fiber Coupled Optical Power Ratio	R_{REF}	0.0125	0.025	0.05	A/W
Ratio of Peak Etalon Photocurrent to Reference Photocurrent	R_{RATIO}	-1.9	0	1	dB
Filter Slope at ITU Lock Point	S_{LOCK}	2.5	3	4	%/GHz
Photodiode Dark Current	I_{DARK}			80	nA
Photodiode Dynamic Resistance			780		Ω
Photodiode Capacitance	C_{PD}		150		pF

Fiber

Parameter	Symbol	Min	Typ	Max	Unit
Fiber Length	L	0.95		1.05	m
Polarization Extinction Ratio	PER	20			dB
Module Optical Isolation		35			dB

Note: Pigtail is single-mode polarization maintaining fiber with 900 µm loose buffer. Standard connector is ST ferrule. Other connectors and fiber lengths may be available upon request.

Thermal

Parameter	Symbol	Min	Typ	Max	Unit
Laser TEC Voltage	$V_{TEC\,LASER}$			2.6	V
Laser TEC Current	$I_{TEC\,LASER}$			1.55	A
Locker TEC Voltage	$V_{TECLOCKER}$	-1.02			V
Locker TEC Current	$I_{TECLOCKER}$	-0.69			A
Thermistor Currents	I_{TC}	10		100	µA
Laser Thermistor Resistance	$R_{TH\,LASER}$	9.8	10.0	10.2	kΩ
Laser Thermistor Alpha	α_{LASER}		-4.39		%/°C
Locker Thermistor Resistance	$R_{TH\,LOCKER}$	9.5	10.0	10.5	kΩ
Locker Thermistor Alpha	α_{LOCKER}		-4.39		%/°C

Electrical Absolute Maximum Ratings

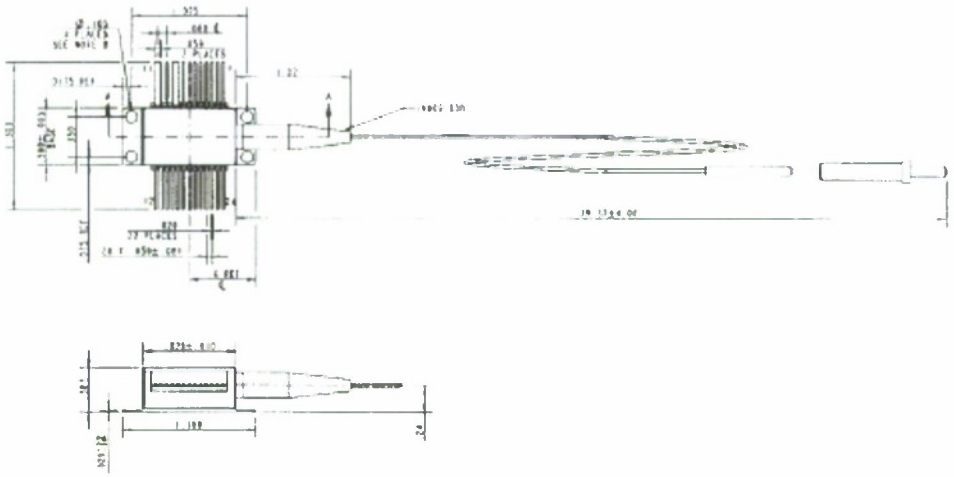
Stresses in excess of the absolute maximum ratings can cause permanent damage to the device. These are absolute stress ratings only. Functional operation of the device is not implied at these or any other conditions in excess of those given in the operations sections of the data sheet. Exposure to absolute maximum ratings for extended periods can adversely affect device reliability.

Parameter	Symbol	Min	Max	Unit
Front Mirror Forward Current	I_{FMAMR}		50	mA
Front Mirror Reverse Voltage	V_{FMAMR}		2	V
Back Mirror Forward Current	I_{BMAMR}		100	mA
Back Mirror Reverse Voltage	V_{BMAMR}		2	V
Gain Section Forward Current	I_{GNAMR}		225	mA
Gain Section Reverse Voltage	V_{RGNAMR}		2	V
Phase Section Forward Current	I_{PHAMR}		30	mA
Phase Section Reverse Voltage	V_{PHAMR}		2	V
SOA Forward Current	I_{SOAAMR}		225	mA
SOA Reverse Voltage	V_{SOAAMR}		2	V
λ Photodiode Forward Current	I_{PDAMR}		1	mA
λ Photodiode Reverse Voltage	V_{RPDAMR}		10	V
Laser TEC Current	$I_{TECLASERAMR}$		2.5	A
Laser TEC Voltage	$V_{TECLASERAMR}$		3.6	V
Locker TEC Current	$I_{TECLOCKERAMR}$		2.5	A
Locker TEC Voltage	$V_{TECLOCKERAMR}$		1.07	V
Operating Case Temperature	T_C	0	70	°C
Storage Case Temperature	T_{stg}	-40	85	°C
Lead Soldering Temperature			250/10	°C/s
Relative Humidity (noncondensing)	RH		85	%

Note: The fiber connector handling/protection (dust) cap and the fiber pigtail materials can deform at temperatures above +85°C.

Mechanical Outline Diagram

Dimensions are in inches.



Patents

This product is protected under U.S. Patent Number 4,896,325 and other patents pending worldwide.

Qualification and Reliability

Agility Communications, Inc. is committed to providing the highest product quality and reliability. Agility is on an aggressive schedule to complete ISO 9001 registration. Additionally, the quality system is developed around a business process model that positions us for migration to the new ISO-9000/2000 standard with no impact to our business.

New product introduction follows a comprehensive gating structure that monitors time, costs, quality and reliability of the product from concept to volume manufacturing, and ensures that products meet all requirements during the life cycle.

The product qualification will fully meet the intent of the Telcordia standards, and qualifies the product in stages of integration, building upon qualified Core Technology Platforms. The laser components will be tested under accelerated life tests using both temperature and current, to determine the MTTF. Hermetic Laser Diode Modules used in conjunction with the appropriate control circuitry will be suitable for qualification to GR-468-CORE.

Laser Safety Information

CAUTION: Class 1M invisible laser radiation when open. Do not view directly with optical instruments.

Invisible Laser Radiation
Do not view directly with optical instruments (magnifiers).
Class 1M Laser Product per IEC/EN 60825-1/A2:2001
40 mW at 1550 nm

CAUTION: Use of controls, adjustments, and procedures other than those specified herein may result in hazardous laser radiation exposure.

The 1XXX versions of widely tunable laser diode modules are classified as Class IM per IEC/EN 60825-1/A2:2001. This product complies with FDA/CDRH, 21 CFR 1040.10 and 1040.11, except for deviations pursuant to Laser Notice No. 50 dated 26 July, 2001.

Viewing the laser output with certain optical instruments (for example, eye loupes, magnifiers, and microscopes) within a distance of 100 mm may pose an eye hazard.

Laser power up to 40 mW at 1550 nm could be accessible if optical connector is open or fiber is broken.

For More Information...

Internet: www.agility.com
Email: info@agility.com
Phone: 805-690-1700
Address: Agility Communications, Inc.
680 Pine Avenue
Santa Barbara, CA 93117

Appendix B: Example Linear Ramp Lookup Table

Appendix C shows an example voltage lookup table which is necessary to create a linear downward ramp. The wavelengths specified are at 20°C with minimal bias on the phase section.

Back Mirror Voltage	Front Mirror Voltage	Output Wavelength (nm)
3.2	2.05	1566.59
3.55	2.15	1566.04
3.8	2.25	1565.61
3.85	2.3	1565.5
4.25	2.45	1564.99
4.55	2.5	1564.59
4.95	2.75	1564
2.3	2.8	1563.57
2.45	3.1	1563.02
2.6	3.3	1562.54
2.7	3.45	1562.05
2.95	3.75	1561.53
3.2	4.2	1561.02
3.35	4.7	1560.51
3.7	2.15	1560.03
4.05	2.25	1559.54
4.3	2.4	1559
2.15	2.5	1558.53
2.25	2.6	1558.08
2.45	2.7	1557.53
2.6	2.9	1557.03
2.75	3.1	1556.55
2.95	3.4	1556.02
3.1	3.8	1555.55
3.45	4.05	1555.02
3.55	4.45	1554.56
4	4.85	1554.03
2	2.15	1553.55
2.15	2.25	1553.06
2.25	2.25	1552.63
2.4	2.55	1552.01
2.55	2.7	1551.56
2.7	2.95	1551.01
2.9	3.05	1550.58
3	3.3	1550.12

3.25	3.65	1549.57
3.55	4	1549.05
3.85	4.3	1548.56
4.2	4.75	1548.03
4.5	2.05	1547.49
4.95	2.15	1547.02
2.4	2.2	1546.62
2.5	2.45	1546.13
2.6	2.6	1545.67
2.8	2.75	1545.16
3.05	2.85	1544.67
3.3	3	1544.14
1.55	3.45	1543.5
3.75	1.55	1543.03
2.05	4.15	1542.5
4.55	1.95	1542
2.3	4.9	1541.49
2.4	4.85	1541.08
2.6	2.25	1540.32
2.7	2.3	1539.91
2.85	2.45	1539.43
3.3	2.8	1538.44
1.65	3.15	1538.03
2	3.25	1537.5
2.1	3.7	1537.01
2.2	4	1536.53
2.3	4.35	1536.07
2.45	4.15	1535.57
2.7	4.95	1534.69

3. LIDAR BEAM STEERING using a MEMS actuator

Table of Contents

Section

Appendices

Introduction

LIDAR is an acronym that stands for **LI**ght **D**etection **A**nd **R**anging. Most people are familiar with RADAR, which stands for **RA**dio **D**etection **A**nd **R**anging and shares many similarities with LIDAR. The idea behind each system is the same and can easily be understood by analyzing a basic pulse, as seen in Figure 1 to the left. First, an electromagnetic wave is emitted, in this case create a pulse. The pulse strikes an object in the distance, causing portions of it to be scattered, others to be absorbed, and some to be reflected back to the transmitter. The signal that makes it back to the transmitter, where a receiver is also located, can be analyzed for various parameters such as transit time, magnitude, frequency shift, and polarization. The transit time allows calculation of distance, based upon the speed of the wave, while the other parameters provide information about the material properties of the object. The main difference between the two systems is that RADAR uses radio waves, which are very broad, and LIDAR uses light waves, which are much more precise.[1]

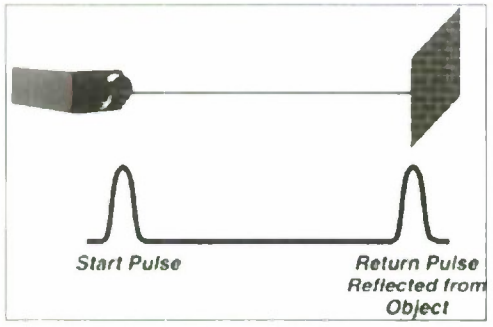


Figure 1: Principle of LIDAR

LIDAR is used for a wide array of applications. As Figure 2 shows, one example is object speed detection. In this system the 'time-of-flight' method is used to iteratively calculate an object's position. The amount of time it takes the light signal to strike the object and return to the operator, coupled

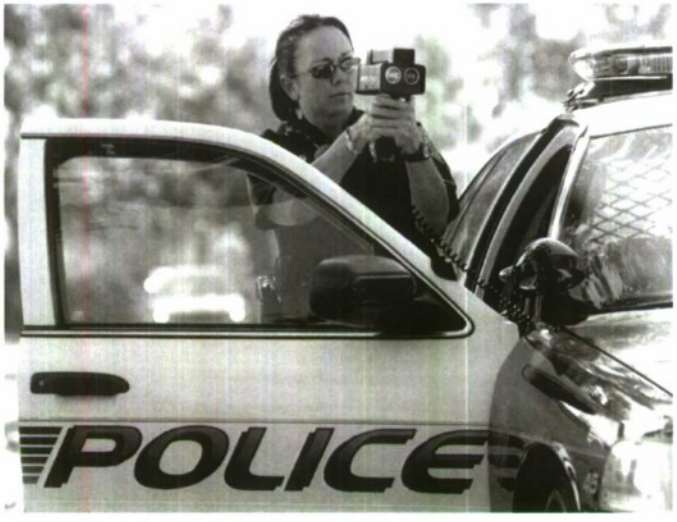


Figure 2: Speed Detection Using LIC

speed of light, allows for a simple distance calculation. Many distance measurements are taken at known time intervals, so the average speed can be quickly and accurately determined.[2]

The most common application of LIDAR, however, is imaging. The two main types of imaging are geographic and atmospheric, each of which use the classic LIDAR analysis techniques described earlier. Geographic imaging is typically carried out by a satellite or airplane, as shown in Figure 3, and the images are generally used to

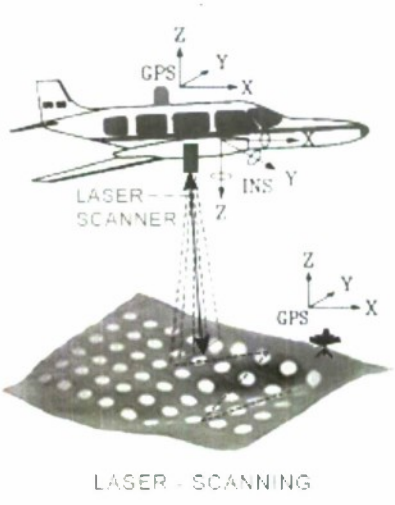


Figure 3: Geographical Scanning LIDAR

Figure 4: Twin Towers Image

generate maps. The intricate detail made available by scanning LIDAR's narrow beam over the area of an object can be seen in Figure 4, which shows the twin towers of the World Trade Center before the 9/11 terrorist bombing. Atmospheric imaging can also be done from an airplane, but the alternative is usually a ground based system. Instead of bouncing light off of the ground, water, buildings, etc. the laser is reflected by aerosols in the air. This practice can be helpful for characterizing wind currents, determining pollution levels, and much more.[3]

Requirements

Dr. Dennis Derickson, Assistant Professor at California Polytechnic State University in San Luis Obispo, had recently been researching single-chip tunable wavelength lasers. These lasers had been used extensively in communications and he realized that many of their strengths would likely provide unique solutions to LIDAR applications, so long as their weaknesses could be sufficiently mitigated. Table I below provides an overview of the strengths of the single-chip laser, while Table II summarized the characteristics which must be suppressed, each of which were borrowed from Dr. Derickson's LIDAR proposal.

Table I: Strengths of Single-Chip Laser to be Embraced

Item	Strength Attribute	Implications of Strength
1	Small size	LIDAR applications become portable and miniaturizable
2	Wide Wavelength Tuning Range	LIDAR with very good distance resolution (on the order of 20 microns)
3	Very Fast Update Rate	LIDAR with microsecond update rates
4	Low Cost	More LIDAR applications can be addressed
5	Adaptable to a wide range of wavelengths	Presently only being designed only at 1550nm. They could be built over the 1000 to 1600 nm range.
6	Low Power Consumption	Enhances portability and miniaturization

Table II: Weaknesses of Single-Chip Laser to be Mitigated

Item	Weakness Attribute	Mitigation Strategy
1	Complicated wavelength tuning mechanism	High-speed driver circuit design effort
2	Single continuous wavelength sweep not available	New system architecture that utilizes a fast stepped-tuned solution
3	Moderate output power	Pursue LIDAR applications that are within the capability of this single chip device
4	Present device packaging is inadequate for LIDAR	Design laser packaging compatible with LIDAR applications

As these tables depict, a single-chip laser would be ideal for a portable, medium power, high speed, and low cost application. Since the imaging sector of LIDAR has more applications and is arguably more interesting, the scope of this project was further narrowed. The task at hand was then to determine which subsystem of LIDAR to approach for design. A common theme across each of the imaging applications was the need to physically scan a laser beam over an area of interest. Thus, the LIDAR Beam Steering project was born. The general requirements to be considered in the design process are outlined in Table III.

	Project Requirements
1	Choose a Mirrored Device to Direct the Laser
2	Develop Control Circuitry for Precisely Controlling the Mirror
3	Provide a Simple Interface for the Single-Chip Laser
4	Instantiate a Modular Design for Further Interfacing
5	Build a Test Fixture to Characterize the Selected Components

Table III: Project Requirements

Design

The first major decision to make in the design process was which type of beam-directing mirror to use. The three types of mirrors investigated were Galvo Motors with mirror attachments, Piezo Mirrors, and MEMS Mirrors. The major grounds for comparison between the three were size, scanning angle, complexity, power, and cost.

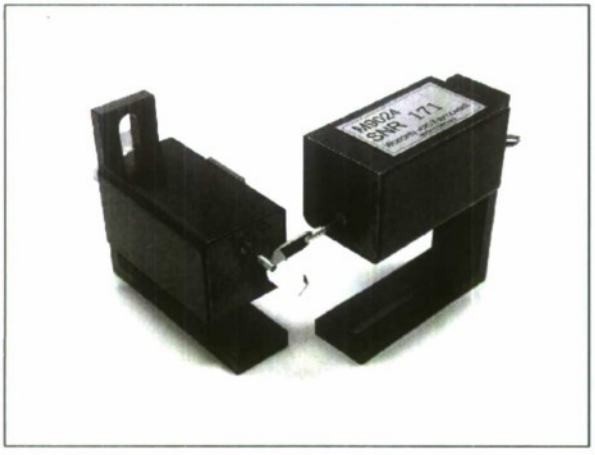


Figure 5: Galvo Motors

A set of Galvo Motors is shown in Figure 5 to the left. These devices can be considered limited rotation electric motors, with the exception that their deflection angle is a function of current (magnitude and polarity). When oriented orthogonal to one another, an incident laser beam can be deflected along two axes. Some of the higher-end galvos use position feedback to produce a positioning system with little to no overshoot, while operating in the tens of kilohertz range. Due to their speed, accuracy, and robust nature, a set of galvos is the preferred choice by laser light show operators.[4] However, their traits do not readily lend themselves to the project at hand. These units are typically several inches in length, each setup requires two motors, and their orthogonal orientation is anything but space conservative. Next, the control software and circuitry required to drive two separate motors with position feedback is rather complicated. Along these same lines, current controlled components require even more circuitry because most driving sources are voltage controlled; high current is also bad because it relates to high power dissipation. Thus, the need for voltage to current converters arises. Lastly, a popular galvo motor set like the General Scanning G120, which comes with both motors mounted

in the appropriate bracket, costs upwards of $1000. So, needless to say, galvo motors were not the ideal choice for this application.

The next beam-steering device considered was the piezo mirror, shown below in Figure 6. Piezo mirrors operate based on the principle of piezo-electric materials. That is, when a voltage is applied to this type of material the crystalline structure deforms predictably. These mirrors are somewhat small (3.5" x 3.5" for the one shown),

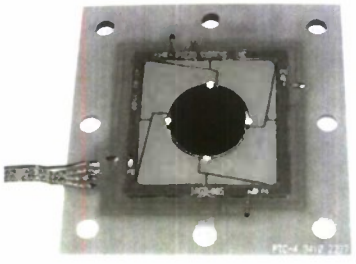


Figure 6: Piezo Mirror

dissipate very little heat, and also have very low power consumption. Only one unit is required to scan in two axes, providing a distinct advantage over the galvo motors. Piezo mirrors can also be used in an open-loop configuration, yielding a significantly less complicated system as well. One of the drawbacks to this type of mirror is that it has a resonant scanning angle of less than two degrees. So, the angle small compared to the galvos, which can scan over entire walls as seen in laser light shows. Finally, the major drawback to the piezo mirrors is the price. To purchase a single mirror with dual-axis scanning capabilities someone would have to spend about $3000. Even though this device appeared promising, the lack of a bottomless budget makes it extremely impractical.[5]

The final beam steering device considered was the **M**icro**E**lectro**M**echanical **S**ystem (MEMS) mirror. These devices can be based on various principles, but are often

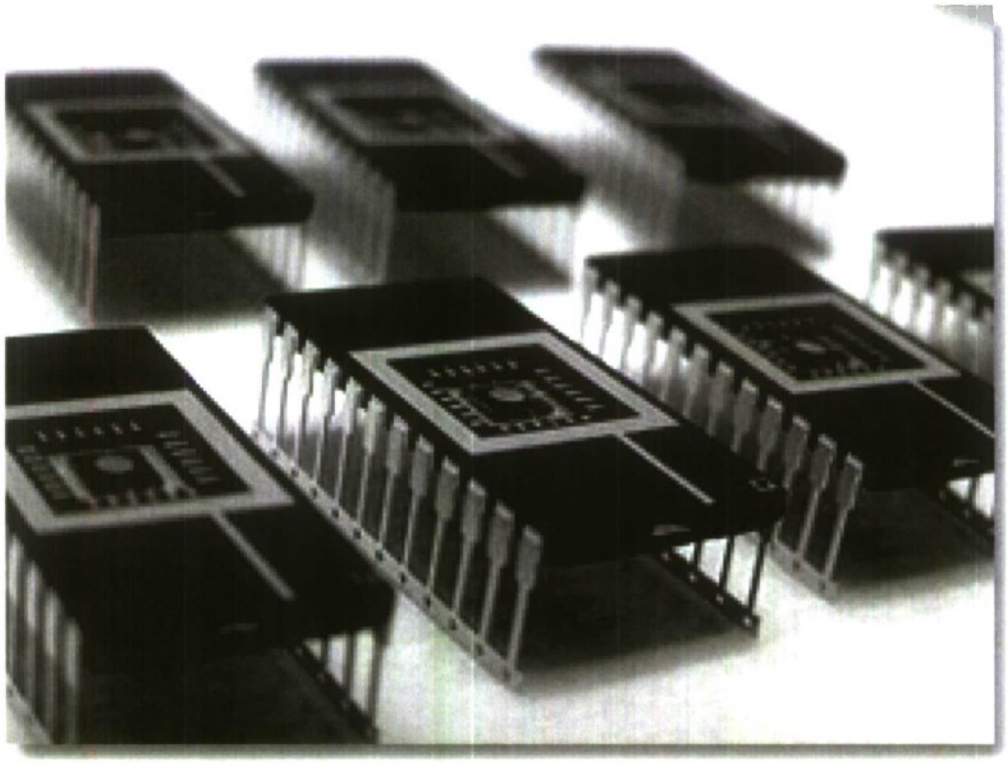


Figure 7: Mirrorcle MEMS On-Chip Mirror

undisclosed due to the infancy of the technology. Figure 7 on the left shows the Mirrorcle dual-axis scanning mirror on a 24-pin pin chip, which likely uses the force generated by a large voltage potential across capacitive plates to

direct its mirror. These devices are

extremely compact, as they can be completely contained within an 8-pin DIP

package. These devices also offer a 10° scanning angle, which is much better than the piezo mirrors. Although the angle is smaller than a galvo setup, it is sufficient to scan medium sized images at a moderate distance, governed by the power limitations of the single-chip laser. Similar to the piezo mirrors, the MEMS mirrors can be operated in an open-loop system, which is again ideal for lower complexity in a new design, and they do not consume a significant amount of power. Amazingly, along with these other ideal traits, the cost is extremely reasonable. A 24-pin MEMS chip with a 0.7mm scanning mirror that rotates from 0 to +10° on both axes only costs about $250. Alas, the MEMS device was the perfect candidate for this application.[6]

Once the chip was selected, the details associated with incorporating it into a system had to be tackled; the specifications for the Mirrorcle device can be found in Appendix D. Since the system was desired to be computer operated, the chip requires voltages up to 120V, and other specialized precautions had to be implemented, the system calls for a **D**ata **A**c**Q**uisition (DAQ) unit to drive the input, a high voltage amplifier for the voltage requirements, and a specialized interface board to handle the additional measures. The block diagram showing these modular components is shown below in Figure 8.

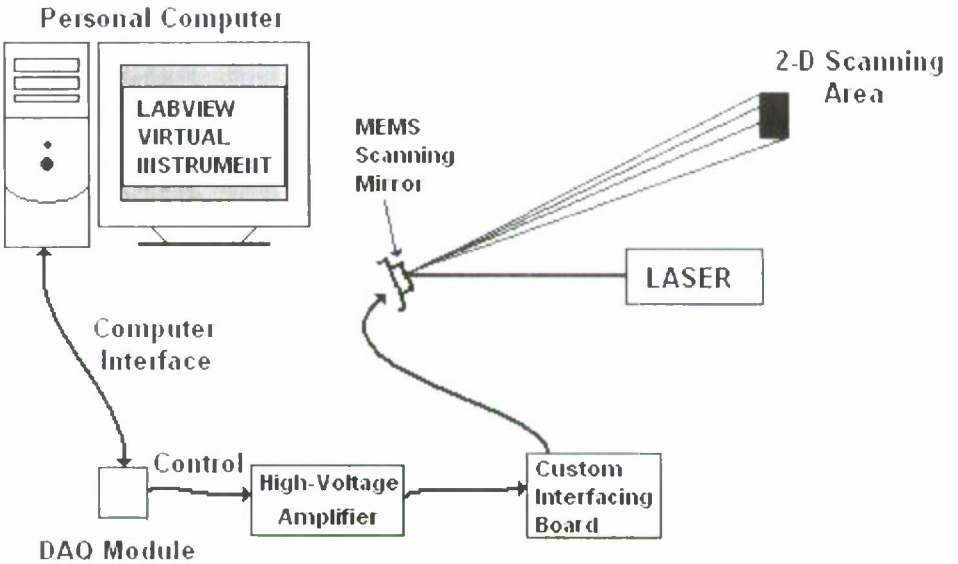


Figure 8: System Block Diagram

Because two axes needed to be controlled independently, a DAQ needed to be chosen with two analog output channels. Also, due to the overwhelming presence of National Instruments LABVIEW software in industry, a National Instruments DAQ was to be selected for its nearly seamless interfacing with the software. Lastly, for simplicity and versatility, a USB connection was decided upon. The DAQ unit that met each of these tasks for a reasonable price was the National Instruments 6008 USB DAQ. Although the output refresh rate was only 150 Hz (samples per second), or a refresh time of 6.67ms between samples, this would be a suitable rate for a first generation system. The datasheet for the DAQ is present in Appendix F.

At the other end of the system, a board was needed to handle the specific needs of the MEMS chip. Namely, a fixture was necessary to physically hold the chip, voltages were not to exceed 120V, and if residual voltage was present after the system was shut down, a means for dissipating this voltage was required. The first solution involved placing a 24-pin chip socket on a prototyping board as shown in Figure 9. To limit the input voltage to the chip, each of the x and y axes received a 120V zener diode in series with a 51kΩ resistor to ground. So, if either input voltage exceeds 120V, the zener breaks down and it is the job of the resistor to limit current flow in the branch. To avoid high currents entering the voltage input pins, precautionary 20kΩ resistors were placed in series with the input to pin connection. Lastly, if the voltages applied to the chip are non-zero when the power is shut-off, the capacitive nature of the inputs store residual charge. To slowly dissipate this, a button in series with a 1MΩ resistor to ground was also placed on each input. The resulting interface-board schematic is shown in Figure 27 of Appendix A. The component values for this schematic are shown below it in Table X of Appendix A.

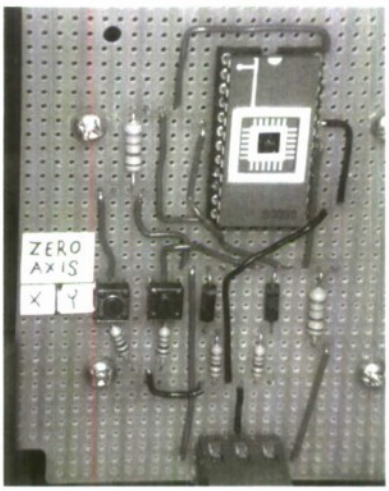


Figure 9: Interface Board

Since each output of the output of the DAQ can provide 0 to 5V output and the mirror

requires 0 to 120V input, a high-voltage amplified was necessary. To avoid driving the DAQ at its limit, the preliminary amplifier design was to map 4V DAQ output to 120V chip input, resulting in an amplifier with gain of 30, or 29.5dB. Also, an upper rail of around 130V would be required to generate an output of 120V. The block diagram for this proposed amplifier setup is shown below in Figure 10. However,

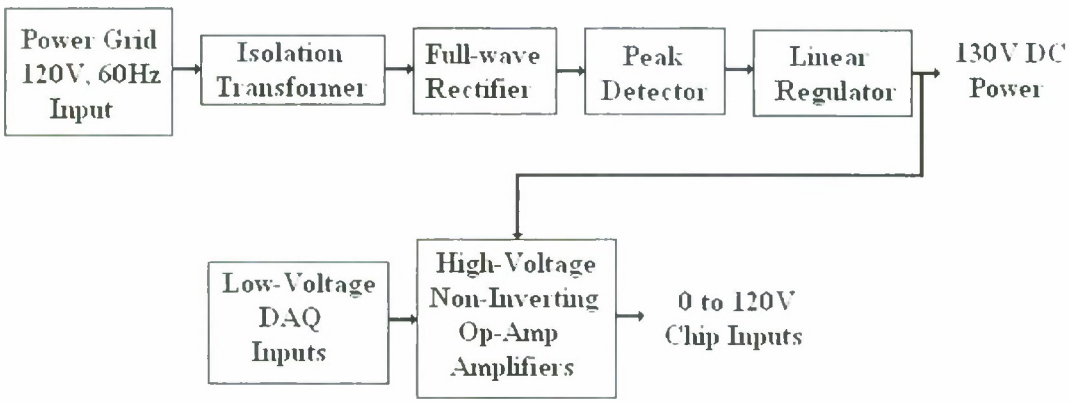


Figure 10: Preliminary High-Voltage Amplifier Design

when the parts cost of this design was compared with an already available amplifier from Mirrorcle Technologies, the appeal of the design strongly regressed. The Mirrorcle HVA032, shown in Figure 11, provided a simple +5V input power connection, which was directly available from the DAQ, eliminating the high-voltage DC generation circuitry. Also, the amplifier had a gain of 60 (35.6dB); so, a 120V output would only require a 2V input voltage from the DAQ. With 12-bits of

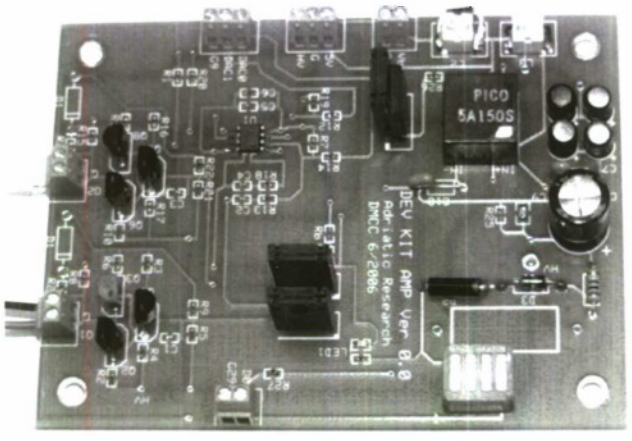


Figure 11: Mirrorcle High-Voltage Amplifier

resolution on each DAQ output channel, a minimum voltage resolution is 1.2mV before the amplifier and 73.2mV on the chip input. So, the low input voltage necessary to provide full-scale output voltage does not present a resolution problem. The datasheet included with the high-voltage amplifier is attached in Appendix E.

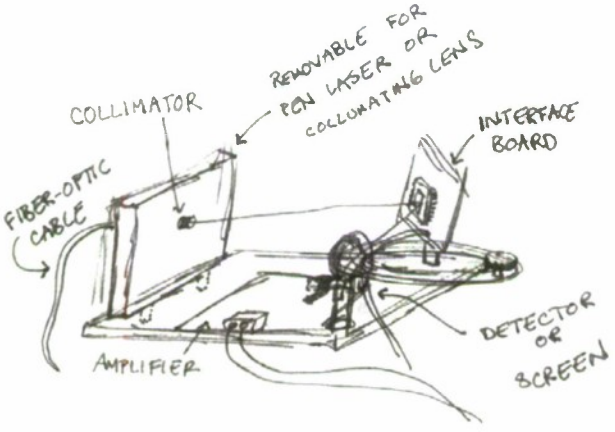


Figure 12: Rough Physical Layout

With nearly all of the physical components selected, ordered, and stored, it was time to design the physical unit which would fasten them in place and allow fixed interconnection. The rough sketch, shown in Figure 12 above, was used as a reference while fabricating the

fixture. The oval shown below the interface board represents a turning mechanism to alter the incident angle of the laser light on the mirror. This feature was added to that future tests can be done to characterize the polarization of reflected light vs. incident angle. The fiber-optic cable allows for a simple interface with the single-chip laser, while the collimator creates a focused beam out of what would otherwise be dispersed light. One of the main additions not shown in this sketch was an arced slot for the power meter/viewing screen to travel in so that it could be quickly moved to account for a change in reflected beam direction.

The final addition to the project was the control software. As mentioned previously, LABVIEW was used to create an interface called a Virtual Instrument, or VI. The VI is composed of two portions, the front panel and the block diagram; these can be seen in Figure 28 and Figure 29 of Appendix C, respectively. The front panel is intended to resemble a physical piece of machinery with knobs, switches, indicators, etc., while the block diagram is where the program structure is developed. The resulting functionality of the VI includes selectable output sample generation timing (≥7ms/sample), numeric displays of DAQ output voltage for each axis, numeric displays which approximate the corresponding voltage produced by the amplifier, graphs to display DAQ output voltage vs. sample number, and a switch to implement a single scan or to scan continuously. This functionality was enabled by first creating two output voltage tasks. The portion of each task that accepts data to be output as a voltage was placed within a while loop, which includes an index that

increments each loop iteration, and a stop button to halt the program. At this point the number of scanning points had to be determined. Since the fastest output sample interval was 7ms, a 64 x 64 array should take approximately 29 seconds. Increasing the scanning resolution to 128 x 128 array would yield a scanning time of 1minute 55 seconds, which is much too long to wait for a single scan, so the 64 x 64 was chosen. With a full tilt angle of 10° per axis, this to 0.156° between each neighboring sample point and a grand total of 4096 individual scanning positions. To create a linear scanning method, triangle waves were generated as outputs. In LABVIEW, this corresponded to creating one period of a triangle wave array for each axis and indexing it according to the loop iteration value. Two arrays were created so that rectangular areas could easily be scanned over by changing the triangle wave array, should the need arise. The index for the x-axis was generated by dividing the loop iteration by 128 to yield a remainder. This creates a number that continuously counts from 0 to 63, then back to zero based on the symmetry of the triangle wave; the count is incremented each clock cycle. The y-axis index uses this same technique, except the index fed into the remainder operator comes from an integer division operator. In other words, the loop index is divided by 64 and the integer portion of the result is passed on to the division by 128 remainder operation. The purpose is to create an index that only increments every 64 clock cycles. The last piece of the puzzle was the continuous vs. single scan switch. The loop index was simply compared with 4095, which corresponds to the final scanning position. If they are equal, a true signal is generated. The switch is used to control a gate which

lets the signal reach the stop button if in single scan mode, or does not let it pass in continuous mode, preventing the stop from occurring. The end result is a program that is very user friendly and makes implementing a scanning operation a breeze.

Construction

All of the parts obtained for the fabrication of this project are outlined in Table XI of Appendix B. The first step was to place the rotational mount, intended for vertical use, horizontally on the base plate as depicted in Figure 13 on the left. This was done by spacing the mount off of the base with the rectangular rod, which was cut into three small pieces. Holes were drilled through the rotational mount and spacers for later mounting.

Figure 13: Components Being Mounting

Next, a hole was drilled in the base plate where the center of the rotational mount was located; this served as a pivot point for the power meter slot. The base was placed in a mill, fastened through this pivot hole, and an arc was created by rotating the base about the hole. The arc was made at a radius just large enough to place

the amplifier between it and the rotational mount. ¼" thick nylon spacers were

purchased so that the amplifier would sit above the base plate when mounted, eliminating short circuits. Then, the post to hold the collimator was mounted within the radius of the power meter slot. It was placed close to the edge of the board in order to maximize the possible incident/reflected angle measurements. With the layout complete, all of the components were placed on the base plate. In the case of the amplifier and rotational mount, holes were drill (as seen in Figure 14 to the left) and threaded for sturdy mounting. Another important piece of the fixture was the interface-board mounting bracket. This was made from a cut-off piece of the base plate and bent at 90° to allow vertical mounting. Slots were milled into the base of

this bracket where it attached to the rotational mount to allow fine adjustment for centering the MEMS chip (forward and back) over the rotational axis of the rotational mount. When the MEMS chip was directly over the center, the mirror's position would be adequate for the laser to strike it at any incident angle without laser repositioning. Lastly, a circular piece of aluminum was cut to be placed in the collimator mounting bracket since the original hole was much too big. Then, a hole only a few thousandths of an inch larger than the collimator diameter was drilled in the center of the aluminum circle for the collimator to be inserted into. To keep the base off of the ground, several non-conductive feet were added. The finished layout is shown below in Figure 15 and annotated for clarity.

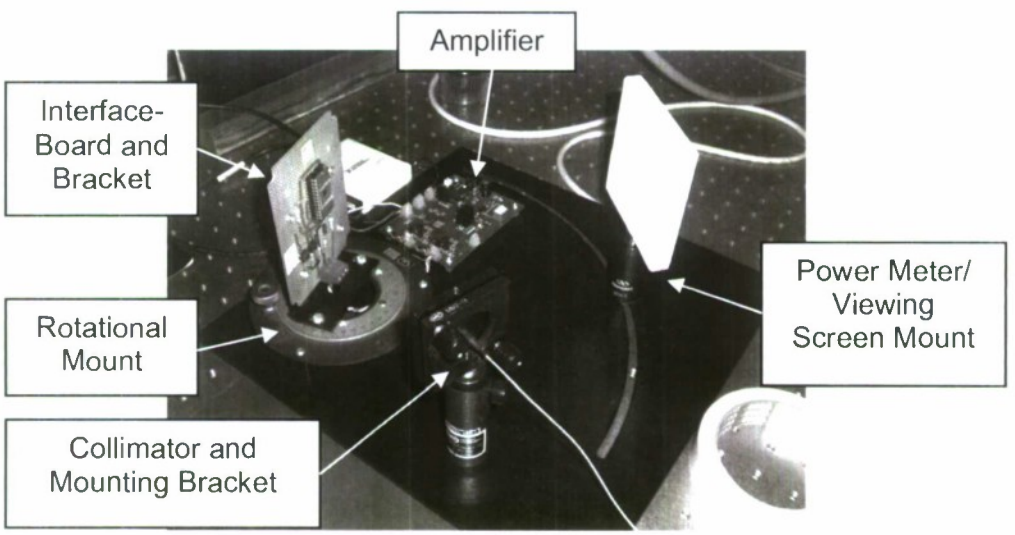


Figure 15: Final System Assembly

The interface board itself had to be fabricated as well. The prototype board purchased for this application was rather difficult to use because only one trace

connected a set of two holes. So, soldering to a node required first connecting several pairs of holes together. The wires for the interface-board were color coordinated (x-axis → red, y-axis → green, ground → black), hand bent, and soldered in place for strength and quality electrical connection.

Wiring the components together was done according to Figure 26 of Appendix A.

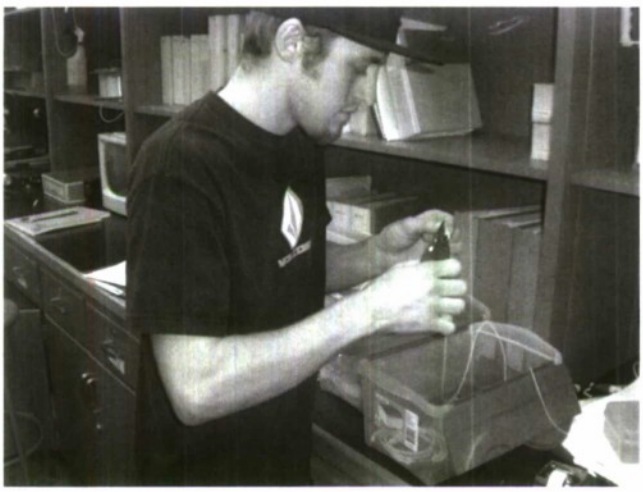


Figure 16: Fusion Splicing

In order to proceed to the testing phase of the project, light had to be channeled into the optical fiber. The solution was to fusion splice another fiber with the correct connector to the end of it. The fusion splicing process involved stripping two outer layers of protective casing, and one layer of protective coating using extremely sharp strippers. Approximately 3" of bare fiber was exposed, cleaned with isopropyl alcohol, and the tip was cut clean with a sharp diamond cutting tool. Each of the two ends received the same treatment, then were placed in the fusion splicing machine. Automated video alignment and a quick, yet strong electric charge fixed the ends together with a low power loss.

Testing

Much of the testing took place in the senior project lab. The first item tested was the interface-board, which was checked for continuity from connector to chip socket. Then, as shown in Figure 17 to the left, the amplifier was characterized while being powered by the DAQ and having no load.

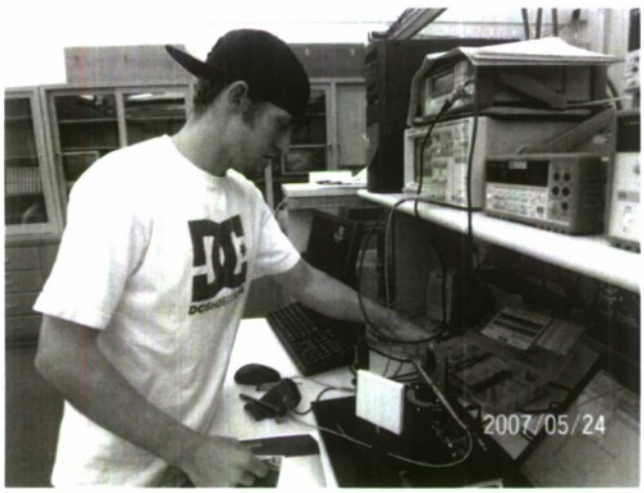


Figure 17: Amplifier Testing

The input to the amplifier was generated using a dual-variable voltage supply and the output voltage measured on a digital multimeter. Table IV on the next page shows both the x and y-axis DC transfer characteristic data. Since a maximum output voltage of 120V was desired, the input to the amplifier should be less than 2V, as seen by the bold row in the table.

V_{in} (V)	$V_{out,y}$ (V)	$V_{out,x}$ (V)	DC Gain	DC Gain (dB)
0.0	0.00	0.00	0.00	0.00
0.1	6.32	6.32	63.20	36.01
0.2	12.45	12.43	62.25	35.88
0.3	18.56	18.52	61.87	35.83
0.4	24.66	24.59	61.65	35.80
0.5	30.74	30.66	61.48	35.77
0.6	36.82	36.72	61.37	35.76
0.7	42.89	42.76	61.27	35.75
0.8	48.96	48.81	61.20	35.74
0.9	55.01	54.85	61.12	35.72
1.0	61.08	60.88	61.08	35.72
1.1	67.12	66.93	61.02	35.71
1.2	73.18	72.96	60.98	35.70
1.3	79.25	78.98	60.96	35.70
1.4	85.29	85.02	60.92	35.70
1.5	91.33	91.04	60.89	35.69
1.6	97.40	97.02	60.88	35.69
1.7	103.42	103.04	60.84	35.68
1.8	109.45	109.06	60.81	35.68
1.9	114.73	115.11	60.38	35.62
2.0	**118.90**	**121.10**	**59.45**	**35.48**
2.1	123.18	127.08	58.66	35.37
2.2	127.43	133.10	57.92	35.26
2.3	131.65	139.05	57.24	35.15
2.4	135.93	145.04	56.64	35.06
2.5	139.19	148.76	55.68	34.91
2.6	139.46	148.90	53.64	34.59
2.7	139.51	148.99	51.67	34.26
2.8	139.54	149.13	49.84	33.95
2.9	139.55	149.19	48.12	33.65
3.0	139.56	149.24	46.52	33.35

Table IV: Amplifier DC Transfer Characteristic Data

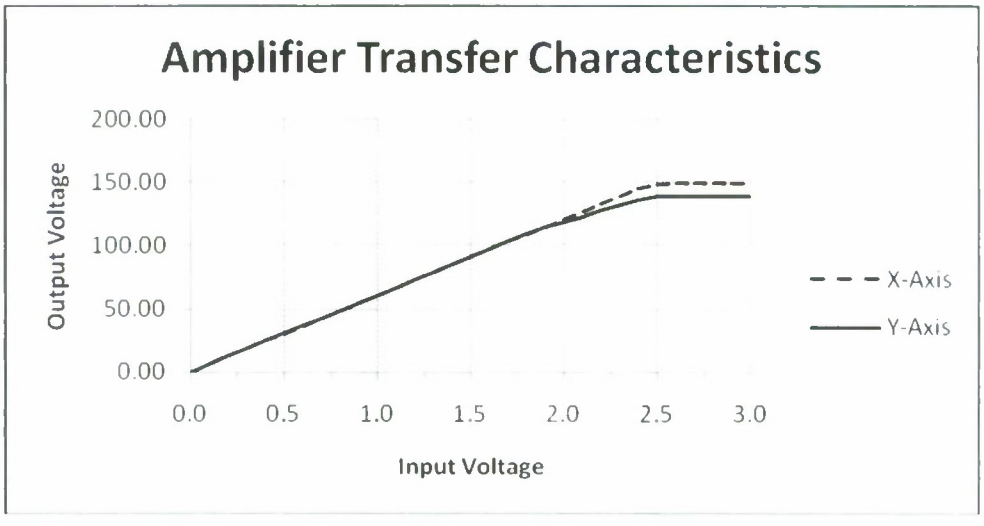


Figure 18: Amplifier DC Transfer Characteristics Plot

The data obtained in Table 17 was used to create the amplifier DC transfer characteristics plot shown in Figure 18, above. As the datasheet alludes to, the maximum voltage output is capped near 150V. The plot is promising because it shows a very linear plot over the intended operating range (to 120V) which can be modeled by the equation in Figure 19 below.

$$V_{out} = \frac{120}{2} V_{in} \rightarrow V_{out} = 60 V_{in}$$

Figure 19: Linear Amplifier Equation

Next, the frequency response data for the amplifier was generated by placing a 0 to

5V sinusoid on the input of the amplifier and measuring the output sinusoid magnitude. The data gathered was used to populate Table V.

Axis	f (Hz)	$V_{in,pp}$	$V_{out,pp}$	Gain	Gain(dB)
Y	50	0.5	30.31	60.62	36
	100	0.5	30.31	60.62	36
	200	0.5	30.31	60.62	36
	500	0.5	30.31	60.62	36
	1000	0.5	30.16	60.32	36
	1500	0.5	29.84	59.68	36
	2000	0.5	29.84	59.68	36
	3000	0.5	29.69	59.38	35
	5000	0.5	29.38	58.76	35
	10000	0.5	28.59	57.18	35
	20000	0.5	26.72	53.44	35
	50000	0.5	18.75	37.5	31
	100000	0.5	10.62	21.24	27
X	50	0.5	30.16	60.32	36
	100	0.5	30.16	60.32	36
	200	0.5	30.16	60.32	36
	500	0.5	30.00	60.00	36
	1000	0.5	30.00	60.00	36
	1500	0.5	29.84	59.68	36
	2000	0.5	29.53	59.06	35
	3000	0.5	29.53	59.06	35
	5000	0.5	29.38	58.76	35
	10000	0.5	28.75	57.50	35
	20000	0.5	27.19	54.38	35
	50000	0.5	20.31	40.62	32
	100000	0.5	12.97	25.94	28

Table V: Amplifier Frequency Response Data
V_{in} = {2.5*cos(2πft) + 2.5]V

The data in Table V was then used to generate the frequency response plot shown below in Figure 20. The 3dB cut-off frequency for each axis is slightly above 30kHz.

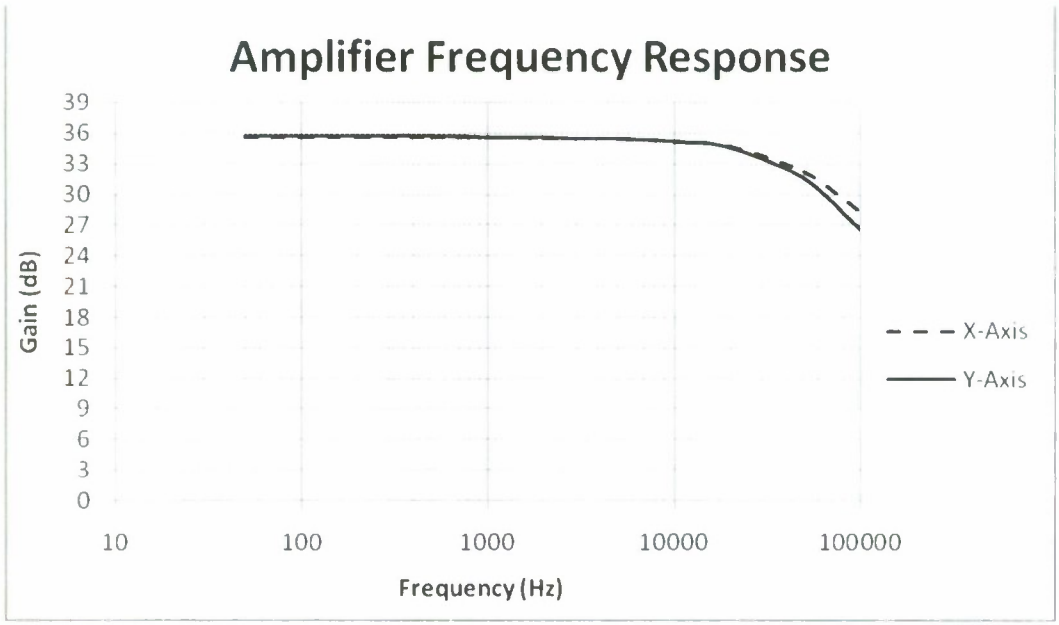


Figure 20: Amplifier Frequency Response Plot

Despite the high cut-off frequency of the amplifier, distortion and phase shifting can be seen at frequencies as low as 9kHz, photographed in Figure 21. As the frequency is increased beyond the amplifier's cut-off frequency, the signal distortion and phase lag continues to get much worse. Figure 22 shows the input and output signals at 100kHz.

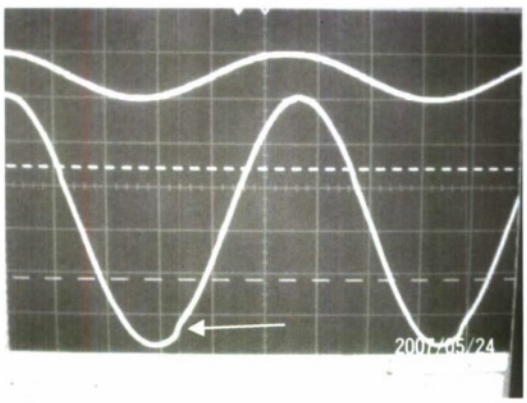


Figure 21: Initial Amplifier Distortion at 9kHz

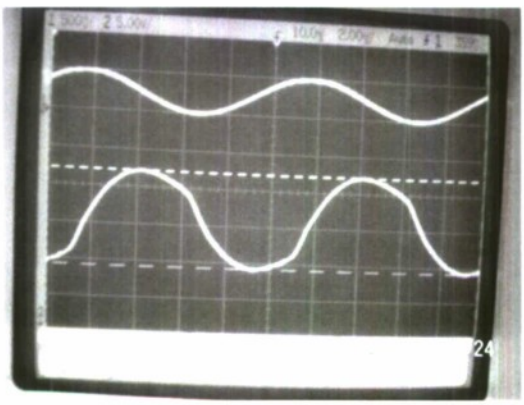


Figure 22: Amplifier Distortion at 100kHz

After much hard work, the system was finally implemented in its entirety. The laser light was reflected by the mirror and deflected from the center position on the viewing screen. For the DC input only a dot was seen deflected from center, whereas an AC signal produced a line from the center to the maximum deflection point.

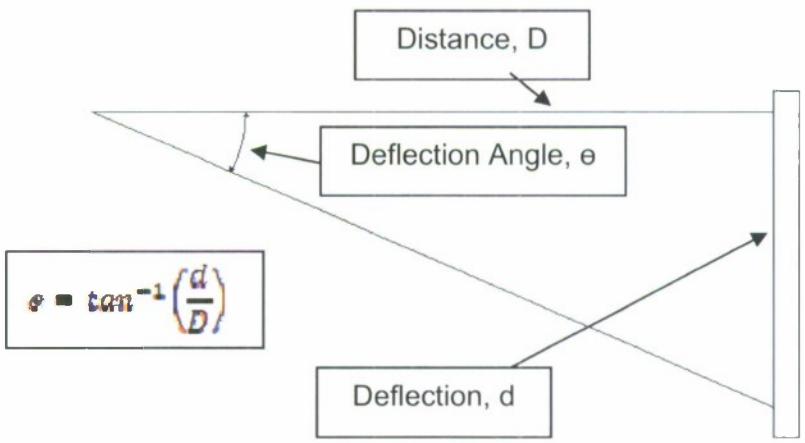


Figure 23: Distance, Deflection, Deflection Angle Relationship

V_{in} (V)	Deflection (cm)	Distance	Angle

0.00	0.00	1.97	0.00
15.65	0.20	1.97	0.29
32.50	0.40	1.97	0.58
46.80	0.80	1.97	1.16
57.00	1.30	1.97	1.89
70.30	1.90	1.97	2.75
80.40	2.50	1.97	3.62
90.50	3.10	1.97	4.47
98.50	3.70	1.97	5.32
110.60	4.60	1.97	6.57
119.40	5.10	1.97	7.26

Table VI: X-Axis System DC Transfer Characteristics

V_{in} (V)	Deflection (cm)	Distance	Angle
0.00	0.00	1.97	0.00
20.87	0.20	1.97	0.29
32.00	0.40	1.97	0.58
48.10	1.00	1.97	1.45
59.80	1.60	1.97	2.32
71.60	2.20	1.97	3.19
80.70	2.80	1.97	4.04
91.80	3.50	1.97	5.04
98.90	4.10	1.97	5.88
107.00	4.70	1.97	6.71
119.40	5.20	1.97	7.39

Table VII: Y-Axis System DC Transfer Characteristics

In either case, the triangle above in Figure 23 was used to calculate the deflection angle based on the deflection distance.

The DC input voltage was supplied via a dual-variable voltage source once again and deflection angle data was recorded in Table VI and Table VII for the x and y-axes respectively. Then this data was used to generate voltage vs. angle plots, as seen in Figure 24 below.

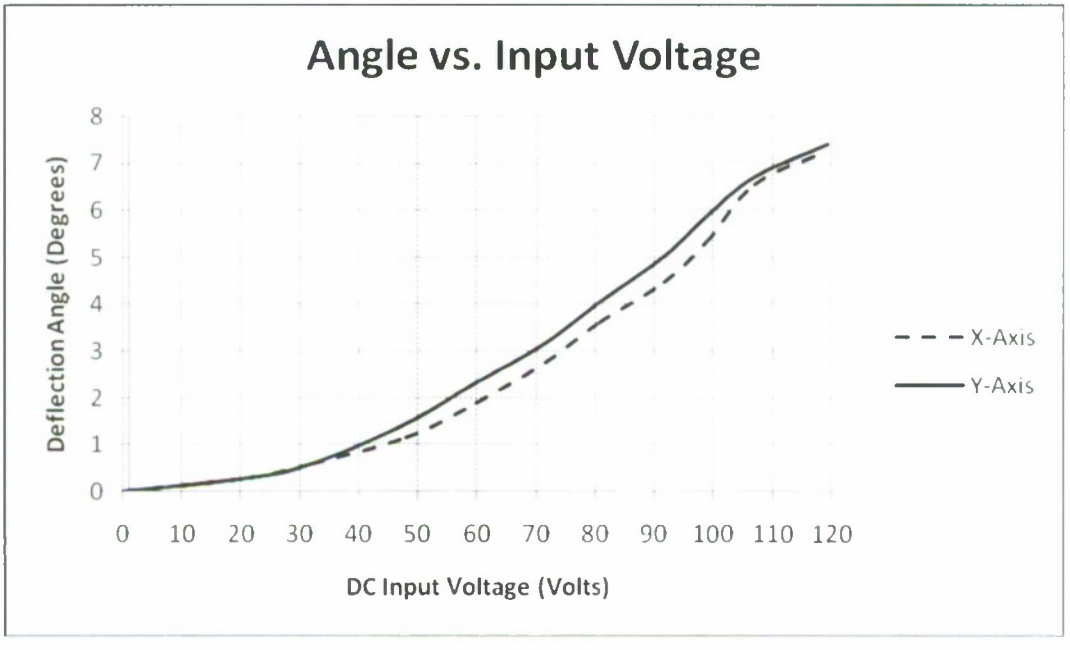


Figure 24: X and Y System DC Transfer Characteristics

For the frequency response measurements a function generator was use to supply

the required input signal while the deflection angle was measured. These results are shown in Table VIII for the x-axis and Table IX for the y-axis. Like the DC characteristics, the AC frequency response for each axis was plotted on the same graph as shown in Figure 25.

f (Hz)	Distance (cm)	Angle (°)
1	0.20	0.58
10	0.20	0.58
100	0.20	0.58
500	0.20	0.58
1000	0.20	0.58
1500	0.20	0.58
2300	0.30	0.88
2890	0.60	1.76
3150	0.90	2.64
3260	1.80	5.27
3350	3.40	9.89
3400	5.10	14.66
3450	7.10	20.01
3490	7.90	22.05
3510	7.30	20.52
3520	2.40	7.02
3600	1.40	4.11
3900	0.50	1.47

Table VIII: X-Axis System Frequency Response Data
$V_{in} = [17*\cos(2\pi ft) + 780]$ mV

f (Hz)	Distance (cm)	Angle (°)
1	0.10	0.29
10	0.10	0.29
100	0.10	0.29
500	0.10	0.29
1000	0.10	0.29
1130	0.50	1.47
1300	0.30	0.88
1500	0.20	0.59
2300	0.40	1.18
2700	1.70	4.98
2780	3.80	11.03
2820	5.20	14.93
2870	6.90	19.49
2900	8.00	22.31
2910	1.40	4.11
2980	0.80	2.35
3090	0.50	1.47
3200	0.30	0.88
3300	0.30	0.88

Table IX: Y-Axis System Frequency Response Data
$V_{in} = [17*\cos(2\pi ft) + 780]$ mV

The measured resonant frequency for the x-axis was 3490Hz, whereas the resonant frequency for the y-axis was measured at only 2900Hz. The values provided in the datasheet are 3691Hz for the x-axis and 3472Hz for the y-axis. This means that there

is only a five percent deviation from expected values for the x-axis, but the y-axis deviates significantly at over 16%.

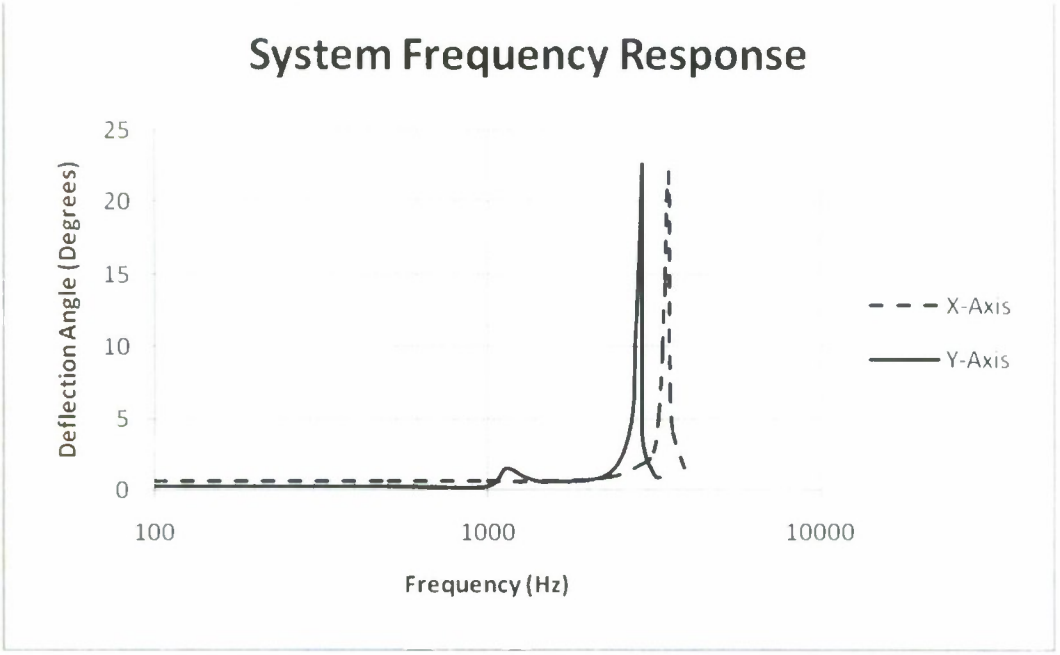


Figure 25: System Frequency Response Plot

During this phase of the testing the y-axis was also responding quite awkwardly. Sometimes with no input the simple touch of a finger on the pins to the MEMS chip would cause the y-axis to stick at a large angle near 5°. Even more strange is that this deflection occurred in the negative y direction, where the mirror is not supposed to be able to travel. This behavior joined with the large error in resonant frequency suggests that the mirror likely has a defective y-axis scanning motion.

Conclusion

This project has produced a working subsystem of LIDAR: in particular, an extremely detailed control algorithm, state of the art MEMS technology scanning mirror, modular test and characterization setup, and all of the interconnections to make it happen. Further improvements to this system would include first returning the chip to Mirrorcle Technologies to investigate the cause of the y-axis sticking; the scanning motion should be smooth and precise. Then, a low-pass filter after the DAQ and before the amplifier would likely be helpful to eliminate EMI noise caused by output discretization, which is essentially a large quantity of step functions. Lastly, as the other aspects of the entire LIDAR system progress, one may consider purchasing a higher output sampling speed DAQ to implement higher resolution images in less time.

It is amazing how much knowledge about the various stages of development has been learned by completing this design from start to finish. Some of the key areas of importance when working on a project of this magnitude, which is small compared to industry projects, are time management, group productivity, punctuality, and most of all enjoying yourself. All of the research, fabrication, testing, and analysis time have added up to over 160 hour of labor, as seen in Table XII of Appendix B, at is has been time well spent. This has been one of the greatest learning experiences of my academic career and these lessons will be carried on into the working world as well as graduate school.

References

1. Dr. Michael J. Kavaya, "LIDAR Tutorial", www.ghcc.msfc.nasa.gov, 2 August 1999, < http://www.ghcc.msfc.nasa.gov/sparclc/sparclc_tutorial.html>
2. "LIDAR Theory of Operation", fortress.wa.gov, Washington State Criminal Justice Training Commission, 4 June 2002, <https://fortress.wa.gov/ cjtc/www/ forms/Lidar_Operator/Lidar_Student_Manual.pdf>
3. "LIDAR Frequently Asked Questions", www.michiganaerospace.com, 2000 Michigan Aerospace Corporation, < http://www.michiganaerospacc.com/business_units/lidar/lidar.shtml>
4. "How Laser Shows Work", www.laserfx.com, 1996-2002, < http://www.laserfx.com/Works/Works3S.html>
5. "2-Axis Piezo Mirror Tilter", www.piezo.com, Piezo Systems Inc., < http://www.piezo.com/prodmirror2.html>
6. "Gimbal-less Two Axis Scanning Micromirrors", www.mirrorcletcch.com, 2005-2007, Mirrorcle Technologies Inc., < http://www.mirrorcletech.com/ pdf/MTI%20Micromirror%20Description.pdf>

Works Cited

Derickson, Dr. Dennis "LIDAR Applications Enabled by Fast Wavelength-Tuning Single-Chip Wavelength Tunable Lasers", Microsoft Word Project Proposal, 2006.

Kavaya, Dr. Michael J. "LIDAR Tutorial", www.ghcc.msfc.nasa.gov, 2 August 1999, < http://www.ghcc.msfc.nasa.gov/sparcle/sparcle_tutorial.html>

"2-Axis Piezo Mirror Tilter", www.piezo.com, Piezo Systems Inc., < http://www.piezo.com/prodmirror2.html>

"Gimbal-less Two Axis Scanning Micromirrors", www.mirrorcletech.com, 2005-2007, Mirrorcle Technologies Inc., < http://www.mirrorcletech.com/pdf/MTI%20Micromirror%20Description.pdf>

"How Laser Shows Work", www.laserfx.com, 1996-2002, < http://www.laserfx.com/Works/Works3S.html>

"LIDAR Frequently Asked Questions", www.michiganaerospace.com, 2000 Michigan Aerospace Corporation, < http://www.michiganaerospace.com/business_units/lidar/lidar.shtml>

"LIDAR Theory of Operation", fortress.wa.gov, Washington State Criminal Justice Training Commission, 4 June 2002, <https://fortress.wa.gov/cjtc/www/forms/Lidar_Operator/Lidar_Student_Manual.pdf>

Appendix A: Schematics

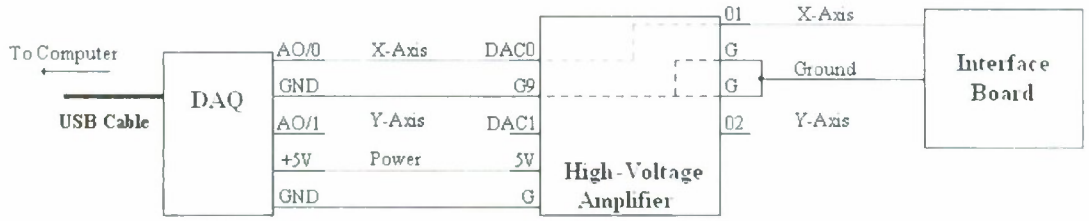


Figure 26: Top-Level Schematic Diagram

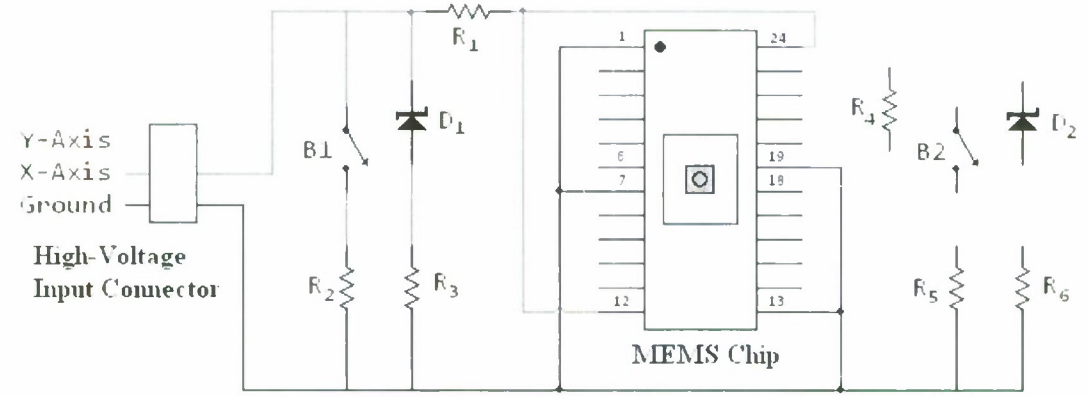


Figure 27: Interface-Board Schematic Diagram

Component	Value
$R_1 = R_4$	20kΩ
$R_2 = R_5$	1MΩ
$R_3 = R_6$	51kΩ
$V_{Z,D1} = V_{Z,D2}$	120V

Table X: Interface-Board Component Values

Appendix B: Parts List, Cost, and Time Schedule Allocation

Description	Cost
National Instruments 6008 DAQ	$160.00
Mirrorcle MEMS Micromirror	$258.50
Mirrorcle HVA032 High-Voltage Amplifier	$275.00
Thorlabs Fiber-Optic Collimator	$57.00
Nuts, Bolts, Spacers, Etc. - 20 → 6/32 x 1/2" - 8 → 1/4" O.D. 1/8" I.D. x 1/4" Nylon Spacers - 6 → 6/32 Jam Nuts - 8 → Sticky Feet - 1 → 1/2" x 4" x 10" Pine	$13.00
Aluminum Stock - 1 → 14" x 18" x 3/16" Plate - 1 → 1/2" x 3/4" x 2" Rectangular Rod	$16.00
Interface Board Components - Pack of Thru-Hole Terminal Connectors - Pack of 51kΩ Resistors - Pack of 1MΩ Resistors - Pack of 20kΩ Resistors - FR4 Prototype Board - Heat Shrink Tubing - 2 → Zener Diodes, V_z = 120V - 2 → 24-pin Chip Sockets - 2 → Push Buttons	$40.00
----------**Total**----------	**$818.50**

Table XI: Parts List and Approximate Cost Breakdown

Quarter	Week	Task Description	Hours
Winter 2007	1/8	Introductions, Meeting Time Allocation	2
	1/15	LIDAR Background Research	6
	1/22	LIDAR Background Research	5
	1/29	Project Definition and Proposal	3
	2/5	Mirror Research	6
	2/12	DAQ Research and Selection, Mirror Selection	5
	2/19	Amplifier Research	8
	2/26	Preliminary Amplifier Design	8
	3/5	Alternate Amplifier Selection	5
	3/12	Component Ordering	2
	3/19	Interface-Board Design and Layout	5
	3/26	Spring Break	
Spring 2007	4/2	Component Procurement, Interface-Board Fabrication and Testing	12
	4/9	Mechanical Fixture Design	5
	4/16	Mechanical Fixture Fabrication	14
	4/23	Amplifier Frequency Response, Amplifier DC Transfer Characteristic	7
	4/30	Fusion Spliced Collimator Fiber, System DC Transfer Characteristics	9
	5/7	System Frequency Response	6
	5/14	LABVIEW Tutorials and Experimentation	15
	5/21	LABVIEW Program Design	17
	5/28	LABVIEW System Integration	5
	6/4	Project Demonstration , Finalization of Write-up	16
---------- **Total Hours Invested** ----------			**161**

Table XII: Approximate Time Allocation Schedule

Appendix C: Program Diagrams

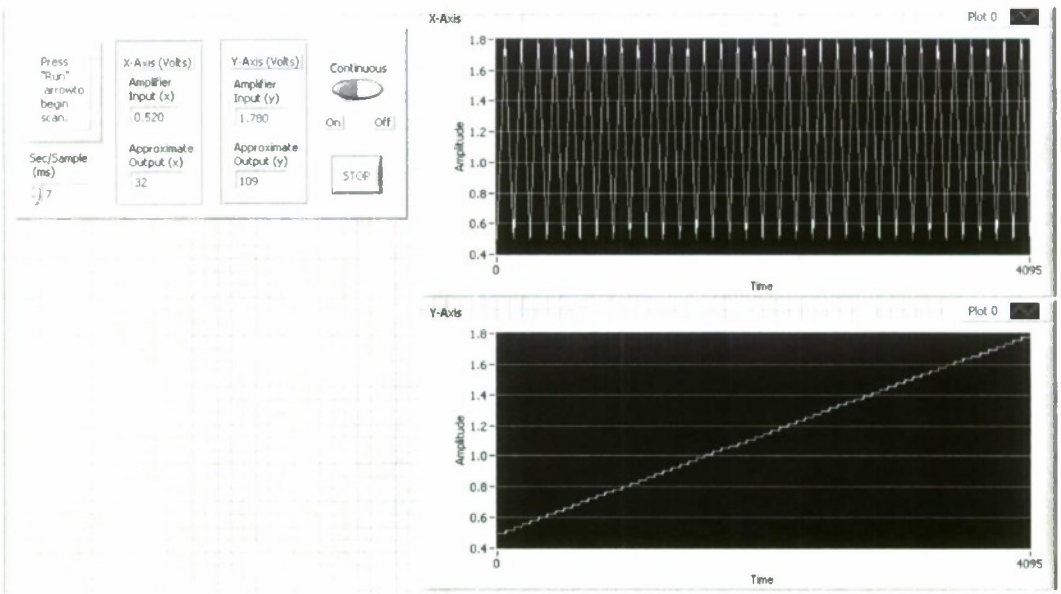


Figure 28: LABVIEW Front Pannell

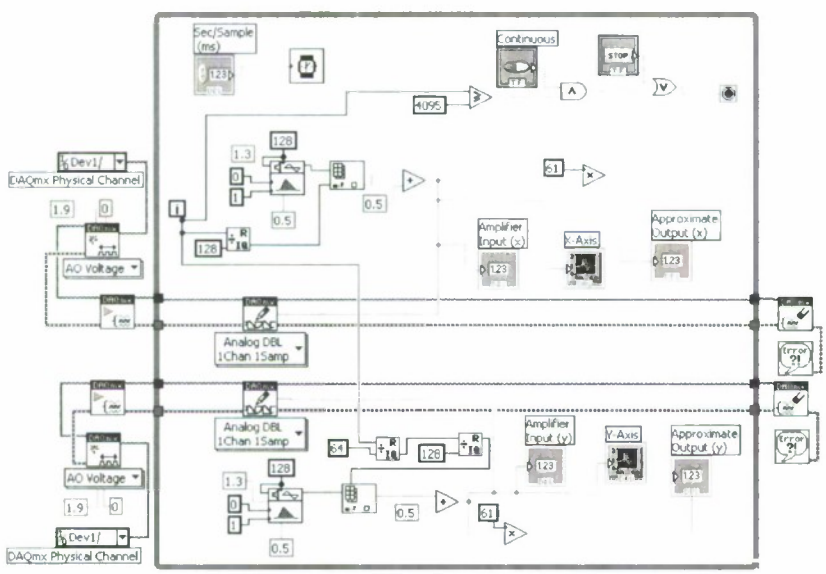


Figure 29: LABVIEW Block Diagram

Shear-Induced One-Dimensional Alignment of Alumina and Silica Nanoparticles in Coatings

Project Investigators:

Raymond H. Fernando, Eric D. Sapper, Lucas J. Brickweg
and Bryce R. Floryancic
Polymers and Coatings Program
Chemistry and Biochemistry
California Polytechnic State University
San Luis Obispo, CA

One-dimensional (1-D) strings of alumina and silica nanoparticles, formed in an extended pearl-necklace fashion were observed near the surface of cured polyurethane films. The nanoparticles have been shown to arrange in perfectly straight lines in the direction of shear. As shear increases, the abundance and length of the arrangements increases. The mechanism of alignment is hypothesized to be a combination of shear-induced patterning and oriented attachment. The alignments have been imaged with atomic force microscopy as well as scanning electron microscopy. Energy dispersive x-ray spectroscopy elemental analysis has confirmed the presence of colloidal nanoparticles in arrangements of smaller particles, while similar arrangements of larger particles that can be seen with an optical microscope have been shown to contain mostly silicon element, presumably a phase-separated polysiloxane containing additive present in all formulations. These findings help advance the knowledge on preparing 1-D aligned particles through template-free approaches.

Introduction

Controlled formation of colloidal particle assemblies has received a great deal of attention in recent years (*1-9*). Applications such as electronics, optics, and medical science could benefit from the unique properties of such highly ordered materials. Some of the previous studies have addressed shear-induced alignment of concentrated colloidal suspensions. Particle size uniformity is an important factor in forming ordered structures under shear flows, and shear-induced particle structure formation has been correlated with shear-thinning at high rates (*1*). The majority of prior research in this area concerns 2-D and 3-D ordered structures formed from highly crowded dispersions.

Literature on 1-D particle arrangements is limited; however, interested readers will find the review by Tang and Kotov (*10*) immensely helpful. 1-D nanoparticle arrangements have been found in nature; for example, the anisotropic structure of maghemite particles in magnetotactic bacteria produces a permanent magnetic dipole, which serves as a critical orientation system (*11*). Particle arrangements of the 1-D type have been prepared by employing numerous methods. Dielectrophoresis and other electrostatic techniques (*6,7,9*) as well as various template-assisted methods (*12*) that make use of entropic force fields to "push" particles in place (*13*) have been reported. While template-assisted methods may yield straight, single nanoparticle wires, a self-assembly method would be favored due to simplicity and compatibility with various integration processes. Previously reported self-assembly methods such as vertical deposition assisted by dewetting have produced well-aligned particle stripe patterns (*14*). Transient particle string formations under shear flows have been reported for highly concentrated dispersions. The transient string formations have been shown to occur as 3-D ordered structures of particles begin to break up, leaving strings of particles aligned parallel to the direction of shear flow (*1*).

This report presents an even simpler self-assembly method for producing stable 1-D nanoparticle arrangements. Shear-induced assembly orders nanoparticles into nearly perfect straight and parallel lines produced by simple application methods.

Experimental

Polyurethane coating films containing either alumina or silica nanoparticles were prepared according to the same methods described elsewhere(*15-17*). Commercially available nanoparticle dispersions (Table 1) were incorporated into the 2K polyurethane clear coat formulation at and below levels of 1.0 wt % (approx. 0.25% by volume) of the cured film. Coating samples were prepared on glass microscope slides using various application methods and cured at 70°C for 30 minutes. Application methods represent various levels of application shear rates and included spray, drawdown, and minimal-shear deposition of liquid coating drops onto the substrate. Final coating thicknesses ranged from 0.25 to 1.5 mils (6-38 μm).

Samples were analyzed with a Pacific Nanotechnology Nano-R atomic force microscope using an Ultrasharp silicon cantilever. Close-contact mode was used for all AFM imaging. A JEOL S-890 scanning electron microscope and a Hitachi S4700 (with Oxford Instruments energy dispersive x-ray spectroscopy (EDS) Detector) scanning electron microscope were used to record electron micrographs and perform elemental analysis.

Results and Discussion

AFM and cross-sectional SEM analysis as well as visual appearance of all the coatings used in this study indicated that, in general, the nanoparticles are well dispersed throughout the film. However, in some occasions, what appeared to be nanoparticles aligned in straight lines were detected. Careful examination of experimental conditions of the occurrences of these 1-D alignments indicated that their formation depends on the coating film application metod.

Application by Drop Deposition Method

Four drops of coating were placed on a glass slide and allowed to flow and spread out for 10 minutes before being cured under the same conditions as stated above (70°C for 30 min.). No linear formations were found by AFM analysis. Very few aggregates of large particles were visible under optical microscope, and aggregates that were present did not display any evidence of linear nanoparticle arrangements when observed with AFM (Figure 1).

Application by Drawdown Method

Drawdown application method using a 75μm-gap drawdown cube produced very straight particle strings in all three of the systems. The alignments are continuous for a few hundred micrometers in many cases. AFM height images of coatings containing Alumina C (0.67 wt.%) and Alumina D (1.0 wt.%) are shown in Figure 2. These samples were prepared at an approximate application speed of 5 cm s^{-1}. Stable, multiple straight-line formations of what was suspected to be alumina and silica nanoparticles were observed. As application speed increased from 1 to 10 cm/sec, the abundance and length of particle strings increased.

Application by Spray Method

Application by spray method produced large strings covering the entire microscope slide surface. The AFM images of these samples show many parallel arrangements in the direction of spray (Figure 3). The AFM line profile shows that most of the smaller particles have heights between 20 and 40 nm, while the larger features have heights over one hundred nanometer range. SEM micrographs of spray-produced

alignments were also recorded, and SEM images obtained at two random locations on a film containing 1.0Wt.% Silica A are shown in Figure 4. Although the modes of deformations in spray applications are complex (*18*), it is a process that involves very high shear rates.

Results presented so far establish the formation of a highly ordered phase when coatings are applied at shear rates exceeding a minimum threshold value. Although the exact threshold shear rate was not determined in this study, estimated shear rates as low as 13 s^{-1} (0.05 cm s^{-1} at 75 μm wet coating thickness) produced strings of particles. However, the apparent sizes of some of the circular features in the images are much too large compared to the known sizes of individual alumina or silica nanoparticles used in the study. In some cases, the 1-D alignments can be seen by an optical microscope at 100X magnification (Figure 5). Surprisingly, we were able to detect few, but definitive 1-D aligned features in films containing no nanoparticles. Contamination was suspected; however, repeated experiments confirmed these patterns are present even when the films were prepared under conditions that eliminated the possibility of contamination by alumina or silica nanoparticles (Figure 6).

SEM Elemental Analysis

As evident from the AFM and SEM images presented so far, many of the features that make up the 1-D strings are much larger than the sizes of alumina and silica nanoparticles. Elemental analysis of the coating film was performed with an SEM equipped with EDS which provided valuable insight into the composition of these features. Electron microscopes operate by firing a beam of electrons or photons at the surfaces of materials. This beam causes inner-shell atomic electrons to dislodge, which eventually create the TEM or SEM micrograph. When the inner-shell electron is emitted a "positive" hole is created, which is quickly occupied by a higher-energy, outer-shell electron. The excess energy from the outer shell electron is given off in the form of an x-ray, which can be measured with an EDS spectrometer. Because every element has a unique atomic and electronic structure, each element emits a different x-ray wavelength when excited by an electron or photon beam. Since the biggest 1-D aligned feature size discrepancy was with spray applied samples, all samples analyzed in this manner were prepared by spray applications on glass substrates. Nanoparticle loading in each case was 1.0 wt.%.

Figure 7 shows EDS analysis results for a spray applied coating sample containing 1.0 wt.% Silica A nanoparticles. The EDS spectrum on the right represents the elemental abundance of a 1 μm diameter area marked "Spectrum 1" in the left image. Presence of Si element is detected. However, when a film containing Alumina D nanoparticles was tested ("Spectrum 1" location in Figure 8, SEM image) not only Al but also Si was detected. When a sample containing Alumina C was analyzed, only Si was detected within a relatively large feature that makes up a 1-D string (see "Spectrum 3" location in Figure 9).

These observations indicate that the shear-induced, 1-D alignments made up of not only the nanoparticles, but also a silicon containing ingredient in the coating formulation. Indeed, siloxane oligomers are commonly used as additives in these types of coatings. When the region of a coating containing 1-D aligned features was subjected to multiple scans within the SEM, craters began to develop (Figure 10). We believe that the emergence of craters in the micrograph is caused by the evaporation of phase-separated silicone due to the heat from the electron beam of the SEM apparatus (Figure 10). Figure 11 contains the SEM image and EDS spectrum of a very small feature after a film containing Alumina C has been subjected to multiple SEM scans. Aluminum element is detected. Similar experiments conducted on films prepared with the formulation containing no nanoparticles showed silicon element within the features that makes up the 1-D strings. In addition, there were other phase separated regions that contained silicon element (Figure 12).

Proposed Mechanism of Alignment

Results presented above establish shear-induced alignment as a template-free method for preparing 1-D strings of alumina and silica nanoparticles. Although their formation mechanism remains somewhat elusive, the siloxane additives in the coating formulations appear to play a critical role. It should be pointed out that, in addition to the siloxane additives incorporated into the formulations, the pre-dispersed nanoparticles (Table 1) contain other siloxane based additives used for stabilizing the dispersions. These

surface active materials have marginal compatibility with relatively polar polyurethane phase. It appears that the shear forces are sufficient to cause separation of the siloxane phase and align in the direction of shear flow. To visualize this on the macroscale, consider a bathroom mirror that has condensed water on it (like a "foggy mirror" after a shower is taken); the water droplets are evenly, yet randomly, distributed on the mirror surface. If a hand is used to wipe away portions of the condensation in a horizontal direction, larger water droplets will align horizontally. This combination of shear and directional application may be the most important factor in forming these nanoparticle alignments. String formation in a sheared polymer blend through a similar mechanism has been reported (*19*).

Also, many properties of the nanomaterials may help provide an intrinsic anisotropy, such as dipole moments of unique axes, anisotropic crystal lattices, and non-uniform stabilizer and surface treatment distributions (*20*). These qualities contribute to Penn and Banfield's crystal growth theory that they have named "oriented attachment". In a typical example of oriented attachment, primary particles align perfectly along one axis, dock, and fuse together under thermal conditions to yield oriented nanocrystal or nanoparticle chains. There is a strong thermodynamic driving force for oriented attachment, as surface energy is reduced substantially when interfaces are eliminated (*20*).

We propose the arrangements presented here arise from a combination of high rates of shear flow, combined with early stages of oriented attachment, possibly aided by uneven distribution of surface treatment on the particle surface. This is similar to ligand-directed assembly of titania nanoparticles with partially removed organic stabilizers (*21*). Furthermore, it is has been shown that oxide nanoparticles are favorably suited for oriented attachment (*22*). Organic ligands, which impede close contact of crystal surfaces, are usually not needed for stabilization.

Conclusions

Atomic force microscopy and scanning electron microscopy equipped with energy dispersive x-ray analysis capability were used to study shear-induced alignment of alumina and silica nanoparticles in two-component polyurethane clear coatings. One-dimensional strings of nanoparticles, formed in an extended pearl-necklace fashion, were observed on both interfaces of the cured films with nanoparticle volume fractions less than 0.05. This alignment is affected by the shear conditions of the application method. When applied by spraying, linear nanoparticle strings as long as 5 centimeters were observed in the direction of shear and application. Linear arrangements of larger, micron-sized particles were present in all coating systems studied, including the control coatings with no added nanoparticle. EDS elemental analysis revealed that these large-particle alignments contain mostly silicon element, believed to be from phase-separated siloxane additive used in all of the formulations. Nanoparticle strings were also found, to a lesser extent, when coatings were applied by a drawdown method. The phenomenon was not observed in coatings applied with minimal shear, such as drop-wise deposition of liquid coating. These nanoparticle alignments, in addition to affecting the performance of coatings, may have broader implications in the field of nanomaterials. Our literature searches have not uncovered reports of stable, 1-D nanoparticle arrangements with equal degree of linearity produced under simple shear with compositions having very low particle loadings.

Acknowledgements

We would like to thank BYK USA, Bayer, and BASF for supplying raw materials, and C^3RP (California Central Coast Research Partnership) and the Office of Naval Research for partial funding related to this project.

References

1) Ackerson, B. *J. Rheol.* **1990**, 34(4) 553-590.
2) Aekerson, B.; Clark, N.A. *Phys. Rev. Lett.* **1981**, 46(2), 123-126.
3) Cohan, I.; Mason, T. G.; Weitz, D. A. *Phys. Rev. Lett.* **2004**, 93(4), 4600-4601.
4) Haw, M. D.; Poon, W.C.; Pusey, P. N. *Phys. Rev. E.* **1998**, 57(6) 6859-6864.
5) Pham, H. H.; Gourevich, I.; Oh, J. K.; Jonkman, J. E. A.; Kumacheva, E. *Adv. Mater.* **2004**, 16(9), 516-520.
6) Kretschmer, R.; Fritzsehe, W. *Langmuir.* **2004**, 20, 11797-11801.
7) Winkleman, A.; Gates, B.; McCarty, L. S.; Whitesides, G.M. *Adv. Mater.* **2005**, 17, 1507-1511.
8) Shenhar, R.; Norsten, T. B.; Rotello, V. M. *Adv. Mater.* **2005**, 17(6), 657-669.
9) Bhatt, K. H.; Velev, O. D. *Langmuir*, **2004**, 20, 467-476.
10) Tang, Z.; Kotov, N. *Adv. Mater.* **2005**, 17(8), 951-962.
11) Dunin-Borkowski, R.E.; McCartney, M.R.; Posfai, M.; Frankel, R.B.; Bazylinski, D.A.; Buseck, P.R. *Eur. J. Mineral.* **2001**, 13, 671-684.
12) Hutchinson, T.O; Liu, Y.-P.; Kiely, C.; Kiely, C.J.; Brust, M. *Adv. Mater.* **2001**, 13(23), 1800-1803.
13) Dinsmore, A.D.; Yodh, A.G. *Langmuir*, **1999**, 15(2), 314-316.
14) Huang, J.; Kim, F.; Tao, A.R.; Connor, S.; Yang, P. *Nature Materials,* **2005**, 4, 896-900.
15) Brickweg, L.J.; Floryaneie, B.R.; Fernando, R.H. *Polymer Materials: Science and Engineering.* **2006**, 51, 124-125.
16) Floryancic, B.R.; Brickweg, L.J.; Comer, J.B.; Sung, L.-P.; Forster, A.M.; Fernando, R.H. *Polymer Materials: Science and Engineering,* **2006**, 51, 251.
17) Brickweg, L.J.; Floryancie, B.R.; Sapper, E.D.; Fernando, R.H. *J. Coat. Tecnhol. Res.* **2007**, 4(1), 107-110.
18) Xing, L.-L.; Glass, J. E.; Fernando, R. H. *J. Coat. Tech.*,**1999**, 71(890), 37.
19) Migler, K. B., *Phys. Rev. Lett.*, **2001**, 86(6), 1023.
20) Alivisatos, A.P., *Science,* **2000**, 289(5480), 736-737.
21) Polleux, J.; Pinna, N.; Antonietti, M.; Niederberger, M. *Adv. Mater.* **2004**, 16(5), 436-439.
22) Pacholski, C.; Kornowski, A.; Weller, H. *Ang. Chem. Int. Ed.*, **2002**, 41(7), 1188-1191.

Table 1. Predispersed Nanoparticles Used in the Study

Particle Type	Dispersion Medium	Solids (Wt.%)	Average Partiele Size
Alumina C	Methoxypropyl acetate	34.0	20 nm
Alumina D	Methoxypropyl acetate	33.0	20 nm
Silica A	Methoxypropyl aeetate/ Methoxypropanol (6:1)	32.0	20 nm

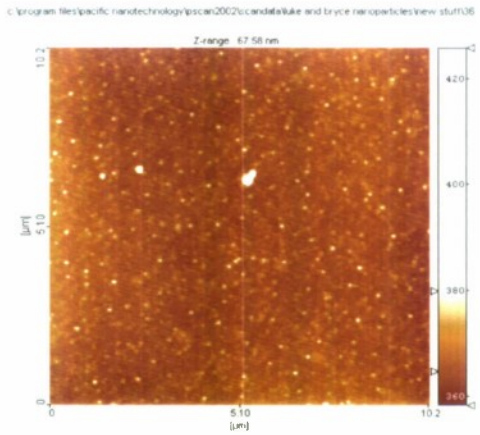


Figure 1. AFM height image (10x10 μm area) of a polyurethane coating containing 1.0 wt.% Alumina C nanoparticles. Application method: drop deposition.

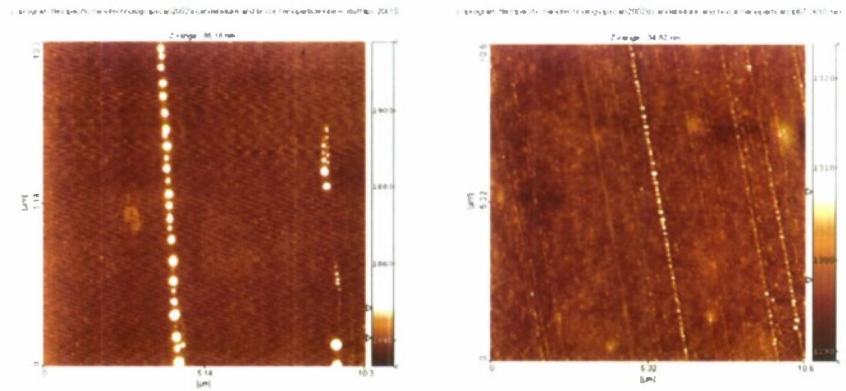


Figure 2. AFM height images (10 x 10 μm) of coatings containing nanoparticles: (left) Alumina-D (1.0 wt.%) and (right)Alumina-C (0.67 wt.%)

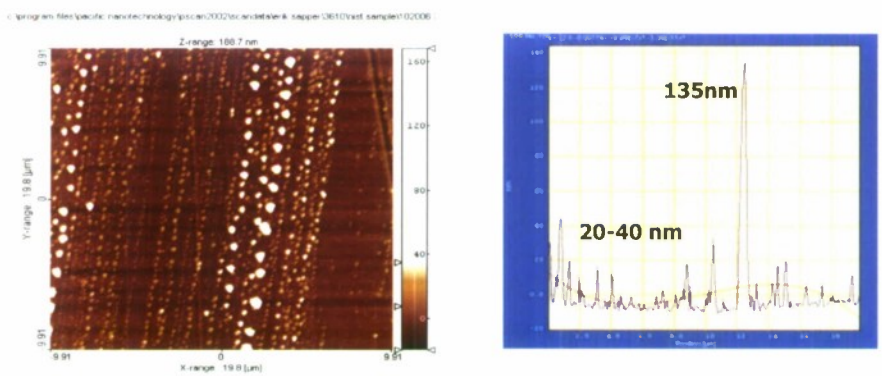


Figure 3. AFM height image (20 x 20 μm) and line profile of Alumina C (1.0 wt.%) arrangements produced by spray application.

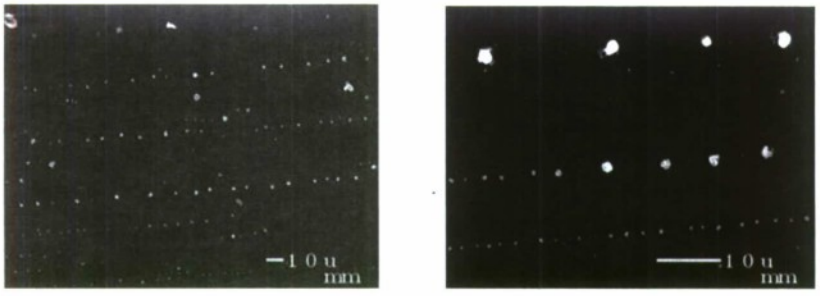


Figure 4. Electron micrographs of Silica A (1.0 wt%) arrangements produced by spray application. The scale bar in each micrograph is 10μm.

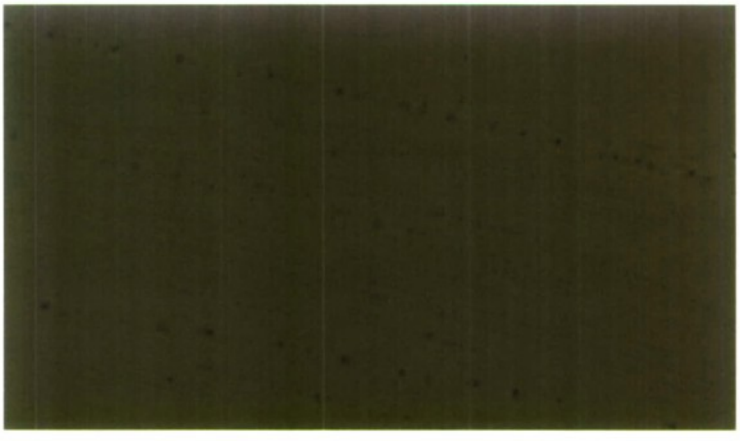


Figure 5. Optical microscopy image of 1-D alignments at 100x magnification. The image represents an approximately 20x12 μm area.

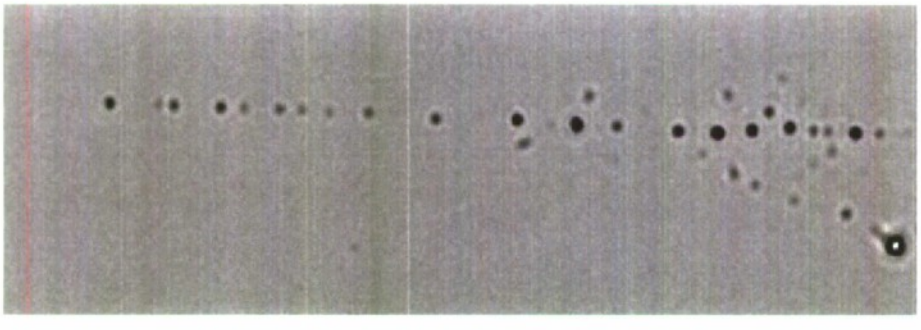
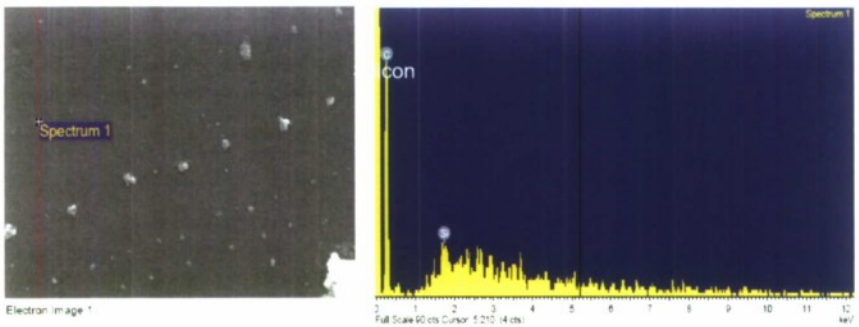


Figure 6. Optical microscope image (400x mag.) of micrometer-scale arrangements produced by spray application of a 2K PU containing no nanoparticles.



Fignre 7. EDS analysis results of a nano-scale aligned feature of a coating containing Silica A nanoparticles.

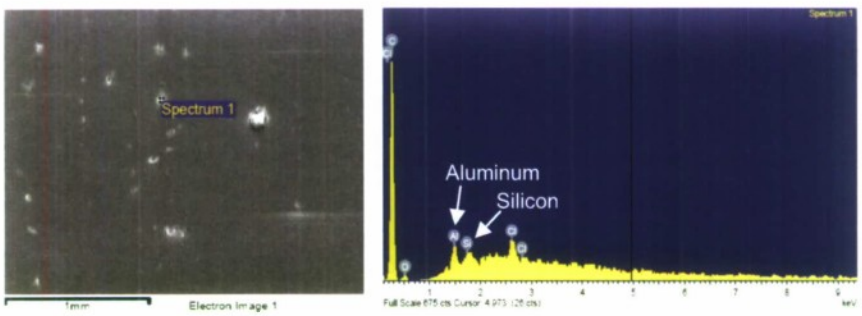
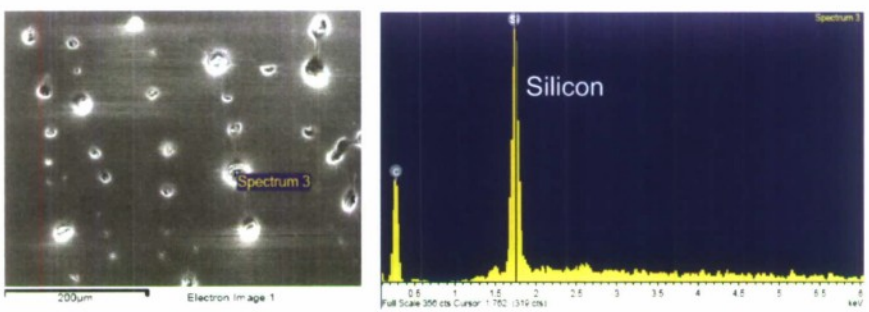


Figure 8. EDS analysis results of aligned particles in a coating containing Alumina D sample shows both Al and Si elements.



Fignre 9. EDS analysis resnlts of larger aligned feature in Alumina C containing coating shows mostly silicon element, believed to be phase-separated siloxane additive in the formulation.

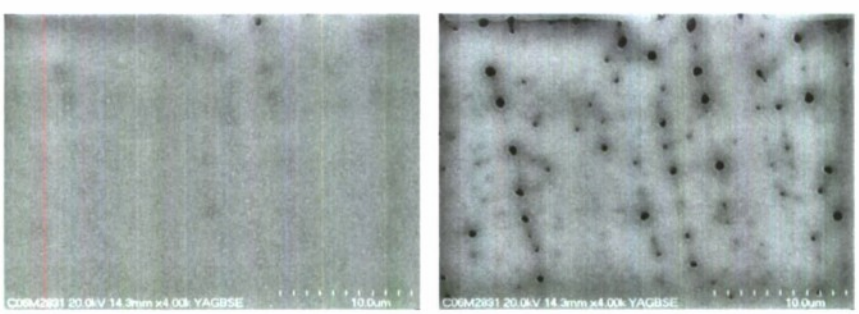


Figure 10. Heat from electron beam of the SEM evaporated large deposits on the coating surface, leaving behind craters on the films surface. The micrograph on the left is after one scan, while the micrograph on the right is the same region after four scans.

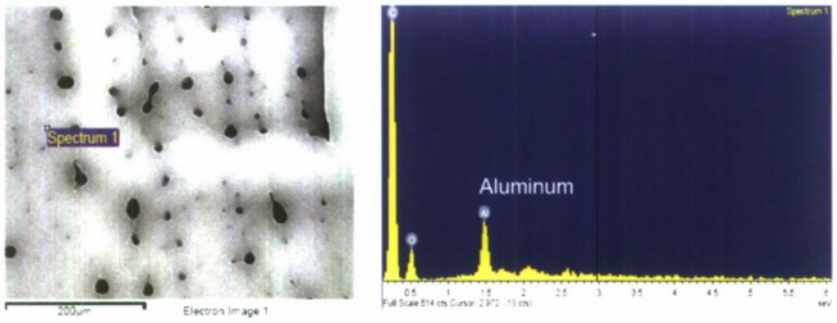


Figure 11. EDS analysis results of Alumina C containing coating shows that smaller aligned particles contain aluminum element, as expected. XRF scans are 1 micron in diameter and are taken in the upper-left area of the SEM image, after four scans

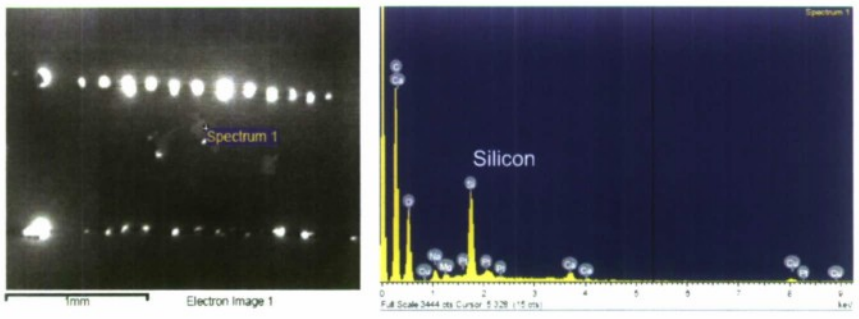


Figure 12. EDS analysis of 1-D arrangements in the control coating system contain primarily carbon and silicon element.

Experimental Investigation of Diffraction Patterns as Atomic Dipole Traps for Quantum Computing

Project Investigator:

Katharina Gillen
Physics
California Polytechnic State University
San Luis Obispo, CA

1. Background Information:

The goal of this project is to assess the diffraction pattern behind a circular aperture for use in quantum computing. Quantum computation and information are currently widely investigated fields because of their applications in communication and decryption, as well as investigation of fundamental quantum physics. One of the systems identified as possible quantum computer are neutral atoms trapped by laser light. One of the remaining difficulties of this system is arranging over one million atoms in a way that still allows addressing individual atoms for quantum computation.

In my previous research I identified the diffraction pattern behind a simple circular aperture as a possible atomic light (or "dipole") trap. Calculations indicated that with modest laser power (~100 mW) atom traps appropriate for quantum computation can be created. Theoretically, an array of circular apertures would then also yield an array of such atom traps.

2. Project Plan:

In this project we performed the next step in the implementation of these proposed atom traps. We measured the intensity pattern behind a commercially available circular pinhole directly to determine whether there are any experimental limitations to the quality of the diffraction pattern for use in quantum computation. For example, one could imagine that the thickness of the pinhole or slight imperfections in circularity may have adverse effects on the diffraction pattern, and thus, the traps.

The experimental setup built was as follows:

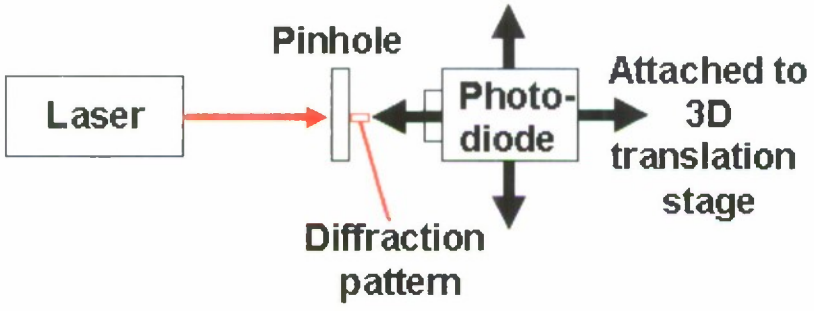


Figure 1: Experimental setup: 100 mW of laser light is incident on a 100 μm diameter circular pinhole. The diffraction pattern behind the aperture is scanned by a photodiode with a 2 μm diameter pinhole, yielding < 2 μm resolution.

We sent 100 mW of laser light onto a circular aperture with a diameter of 100 μm and performed a three-dimensional (3D) scan of the intensity pattern behind the aperture with a photodiode. The photodiode had a sensitivity ~ 1 nW, and a 2 μm diameter pinhole in front of the photodiode yieleded a spatial resolution of < 2 μm.

Using the relationship between the intensity and the trapping potential (shown below), we calculated the trap properties of this pattern for ^{85}Rb atoms, using 780 nm laser light.

The atomic trapping potential is proportional to the light intensity as given by:

$$U(\vec{r}) = \frac{2}{3}\frac{\hbar\Gamma}{8}\frac{\Gamma}{\Delta}\frac{I(\vec{r})}{I_{sat}} \quad (1)$$

Here $\vec{r}$ refers to the position, $\hbar$ is Planck's constant, Γ and I_{sat} are the natural linewidth and saturation intensity of the atomic transition, respectively, and $I(\vec{r})$ and Δ are the intensity and detuning of the trapping laser light.

3. Student involvement:

All aspects of this research project involved undergraduate students. During Spring and Summer quarters four students worked on this project (Angelica Davidson, David Gilbert, Alison Goodsell (Spring only), and Eric Muckley). David Gilbert and Angelica Davidson continued their work on this project in Fall as part of their senior projects. David Gilbert has completed his senior project research, and Angie will continue her work in the Spring of 2008. Funding for this project was used for equipment and supplies, and student salaries.

4. Tasks completed:

The tasks we achieved are:

- Acquire equipment
- Build tunable diode laser with 100 mW output
- Write LabVIEW code to control motion of 3D stage for scanning intensity pattern
- Build and align system: laser, pinhole, 3D stage with photodetector
- Take intensity scans
- Analyze data
- Calculate intensity patterns and trap properties

The equipment needed for this experiment includes lasers, laser power supplies, pinholes, photodetectors, a motorized 3D translation stage (computer-controlled), optics and optical mounts. The lasers were home-built from custom machined parts I brought with me to Cal Poly. The students (Angie Davidson, Alison Goodsell, Eric Muckley) put together every layer of the tunable diode laser, installing the laser diode, collimating lens, and temperature control components. Custom electronics for use in controlling the laser frequency were built by one of the technicians (James Hilsinger) in the Physics department.

David Gilbert put together the photodiode, built the diffraction pattern measurement setup and aligned the laser and optics. In order to fully analyze the diffraction pattern it is necessary to move the photodetector in all three dimensions of space, and record the intensity measurement. To do this, David Gilbert wrote a computer program in LabVIEW which moves the detector along two dimensions (2D), records the intensity measurement at each location, and writes the data to a file for

further analysis. Once a 2D scan is finished, the program moves a step along the longitudinal direction and repeats the process.

The data is analyzed using Igor. David Gilbert developed an Igor procedure that allowed him to extract graphs of the radial intensity pattern at the trap locations as well as graphs along the longitudinal plane. From this data we were able to assess the quality of the actual intensity pattern behind a commercially available circular aperture for quantum computing.

5. Results:

We scanned and analyzed the intensity pattern behind a 100 μm circular aperture. First, we wanted to experimentally confirm the *shape* of the diffraction pattern as predicted by our previous theoretical calculations. Second, we wanted to check whether the trap properties *quantitatively* match our predictions.

As can be seen in the intensity plots below, the shape of the pattern indeed matches our predictions. The diagrams show the calculated pattern in the longitudinal plane (part (a)), as well as our measured patterns in the longitudinal plane (part (b)) and the transverse intensity pattern at the dark spot trap (part (c)) and the bright spot trap (part (d)).

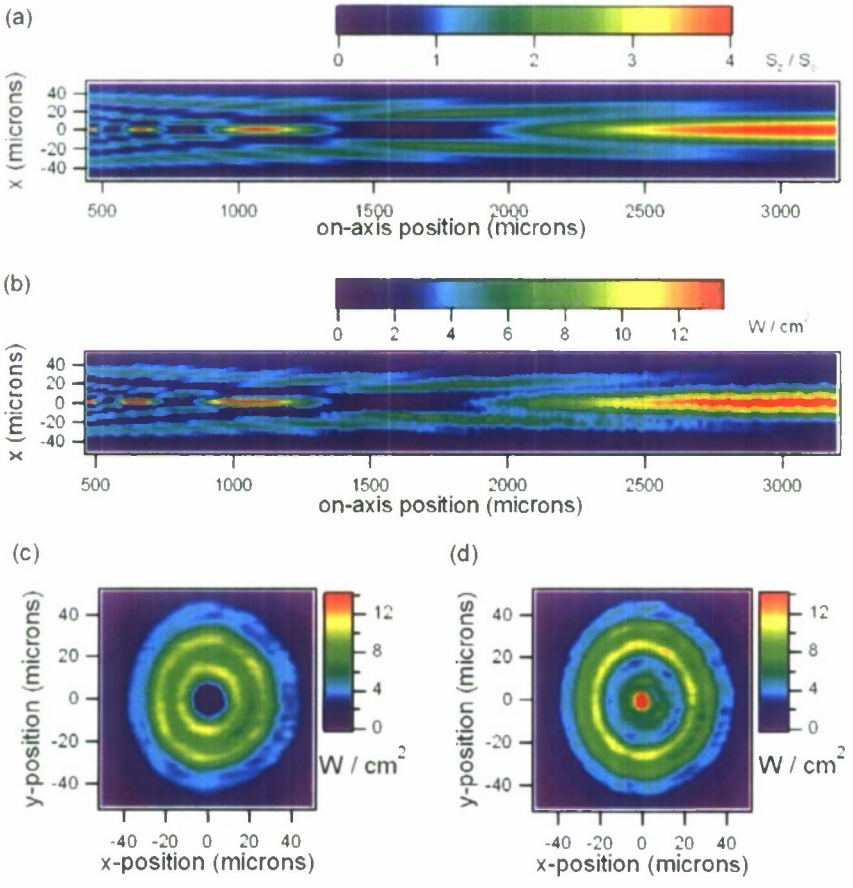


Figure 2: Intensity distribution behind a 100 μm diameter circular aperture. (a) along the longitudinal plane, calculated, normalized. (b) along the longitudinal plane, measured. Note that, due to misalignment at the ~1 micron level, consecutive "slices" of scans had to be aligned by hand, resulting in the visible "roughness" of the diagram. (c) along the transverse plane, at the dark spot trap (1600 μm from the aperture). (d) along the transverse plane, at the bright spot trap (1100 μm from the aperture).

We calculated the trap depth and trap frequencies of our intensity pattern for ^{85}Rb atoms, using 780 nm light.

For the *dark spot trap* (blue-detuned), and a laser detuning of 1,000 natural linewidths, we found:

- Trap depth: 40 μK
- Longitudinal trap frequency: 70 Hz
- Radial trap frequency: 140 Hz

For the *bright spot trap* (red-detuned), and a laser detuning of 10,000 natural linewidths, we found:

- Trap depth: 16.5 μK
- Longitudinal trap frequency: 14 Hz
- Radial trap frequency: 70 Hz

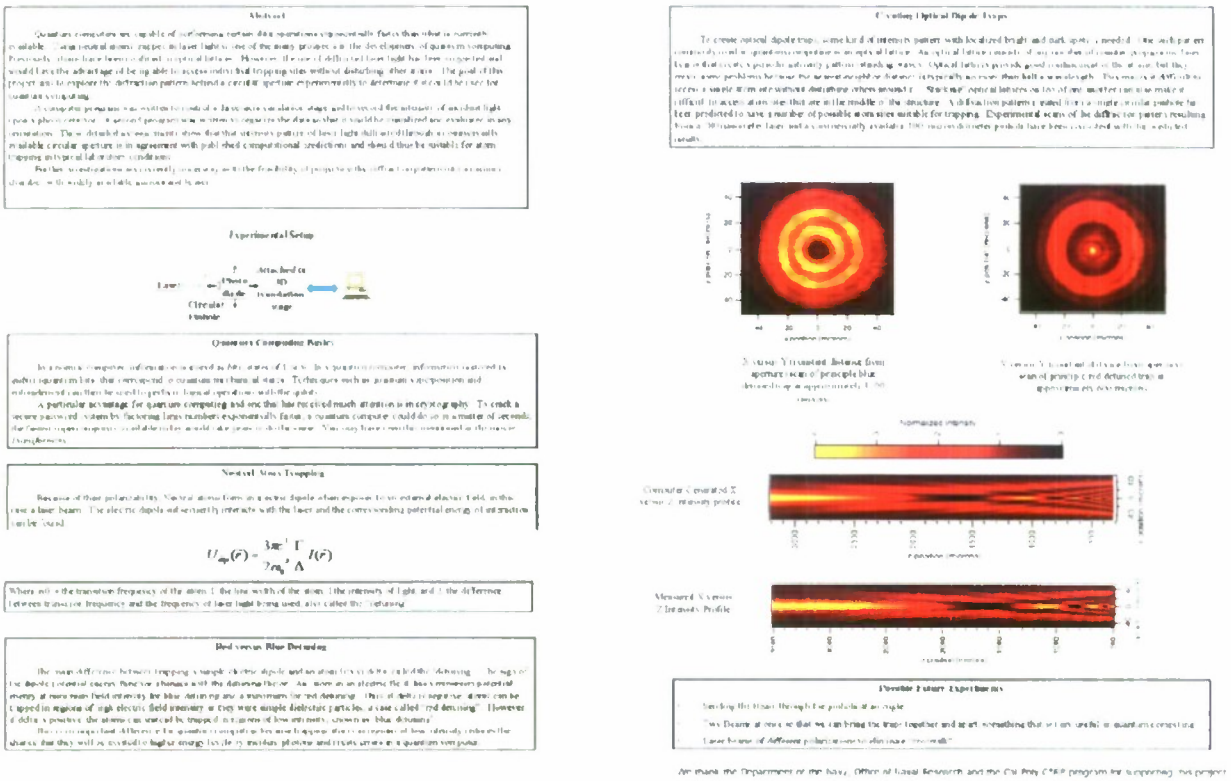
These scans were taken with an estimated laser intensity of 4.5 W/cm^2. This is a factor of 25 (for the blue-detuned trap) and 80 (for the red-detuned trap) less than was used for the calculations, meaning that all the numbers listed above would be 25× or 80× larger for the higher laser intensities – making the experimental results consistent with our theoretical predictions (with the trap frequencies slightly lower, but the same order of magnitude as expected). The higher laser intensities can easily be achieved by focusing the laser beam. We refrained from placing any additional optical components (such as a focusing lens) into our setup for this first scan, so that we could clearly distinguish any effects from the aperture from distortions due to the focusing optics.

6. Presentation of results:

The results of this project were presented by David Gilbert (undergraduate student) at the Southern California Council of Undergraduate Research (SCCUR) conference 2007, at Cal State LA, on November 17th, 2007. The poster is shown below:

Diffracted Laser Light for Optical Dipole Trapping and Quantum Computing

David Gilbert, Physics Major, California Polytechnic State University, San Luis Obispo
Mentor: Katherina Gillen, Department of Physics, California Polytechnic State University, San Luis Obispo



Am ahana the Department of the Navy, Office of Naval Research and the Cal Poly CSUP program for supporting this project

We intend to present these and future results at APS Division of Atomic, Molecular, and Optical Physics (DAMOP), May 27-31, 2008, at Penn State, and to publish the results in the American Journal of Physics or Physical Review A (the choice of journal will depend on the findings).

7. Conclusions:

We successfully measured the intensity pattern behind a commercially available circular aperture and found that the shape of the pattern is as predicted by theory. Quantitatively, the data taken so far indicates that the traps are as good, if not better than what the calculations showed.

In the future we will use this same setup to measure the following diffraction patterns:

- behind a 25 μm aperture
- projected diffraction pattern, in order to determine whether the pattern can be projected into a vacuum chamber
- laser beam at an angle
- two laser beams at an angle
- two laser beams with different light polarizations

This will help design our atom trapping experiment, and determine whether these traps could be used for 2-qubit gates for quantum computing.

All of these measurements can be performed with the same setup and software, and only minor adjustments to the optical system.

Measuring Ocean Surface Velocity and Improving Estimates of Coastal Surface Currents in San Luis Bay with Application to Other Bays and Coastlines

Project Investigator:

A. Elizabeth Griffith
Physics
California Polytechnic State University
San Luis Obispo, CA

This report describes our research project to quantify ocean surface currents in San Luis Obispo Bay using a custom-designed fleet of ocean drifters. Several drifter designs were tested by observing prototype performance at the Cal Poly pier, and a final design was chosen. By the end of the summer a complete fleet of drifters was built and used to collect preliminary data on the surface currents. By characterizing surface currents and discovering mechanisms that drive them, we will be better able to predict coastal surface currents in near-real time. This could help reduce the required time in the water for soldiers swimming to foreign beaches.

A critical requirement for the drifter system was to obtain a 10 mile tracking range from the Cal Poly pier to provide extensive capability for current mapping. This translated into three specific requirements for the real-time tracking system: (A) maintain drifter antennas 4' above the surface of the water, (B) use high power transmission (1 Watt) from each drifter, and (C) mount pier antenna 50' above surface of water to extend the line-of-sight range.

To maintain the stability of the drifter (by counterbalancing the weight of the antenna at the top), the drifter base had to be heavy and low. Experimentally, we found that using a steel bottom ring at a depth of seven feet provided ample stability under typical ocean conditions.

The requirement of seven foot depth meant that the bottom seal of the drifter was exposed to an additional external pressure of about 1/5 of an atmosphere. This caused failures for several material combinations (stand alone plug, plug / silicone paste, plug / aquarium sealant). The best material combination was plug / epoxy, which survived all testing.

We completed several 24-hour float tests using drifter protoypes tethered at the Cal Poly pier. We experimentally verified material and design choices.

In addition, we tested backup systems such as the water-tight electronics cases and ring buoy for secondary flotation.

We conducted reliability tests to check the prolonged effect of seawater on the epoxy/PVC joints.

Plate 1. Prototype testing by Tom Baker and Robin Reil, Cal Poly physics majors.

Plate 2. Overnight Leak Tests at Cal Poly Pier

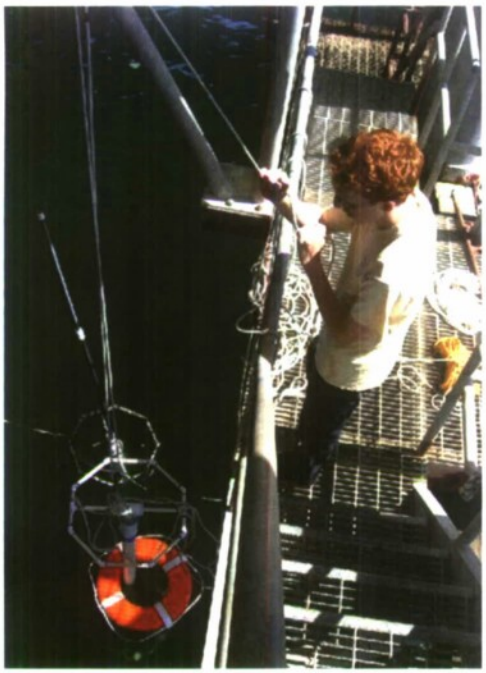
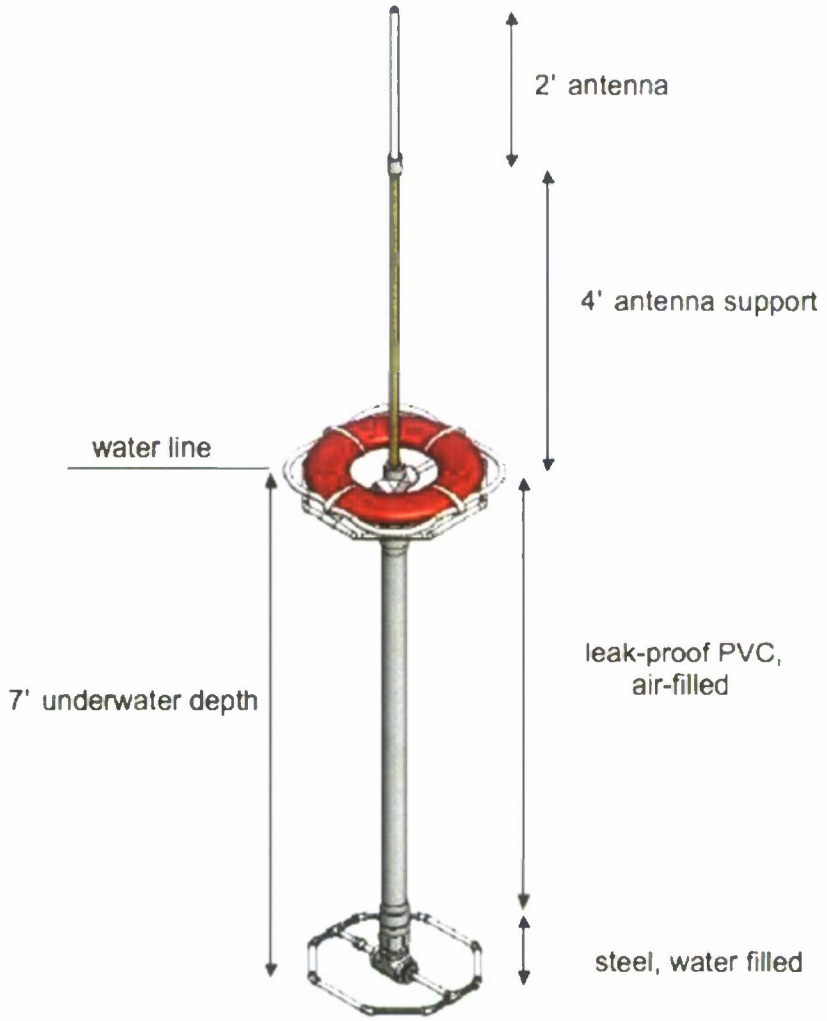


Figure 2. A 3-D Mechanical Drawing of the Final Drifter Design



The software model of the drifters (Figure 1) was created this fall by mechanical engineering undergraduate Nick Supat. The model is a 3-D mechanical drawing that is precisely to-scale. Nick Supat is the third undergraduate to work on the project.

Plate 3. Prototype of final design.

Onee the design was proven, Tom Baker measured, cut, PVC-cemented and epoxied together an additional seven drifters.

Plate 4. Manufacture of drifter fleet.

Plate 5. Drifter-tracking electronics purchased from GPS Flight.

Plate 6. Backup waterproofing system to protect drifter electronics.

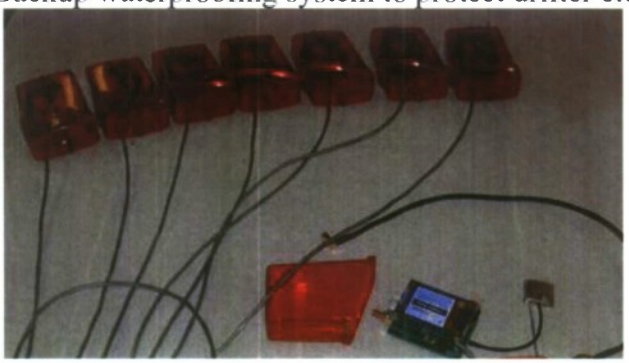


Plate 7. Assembly of first drifter at pier.

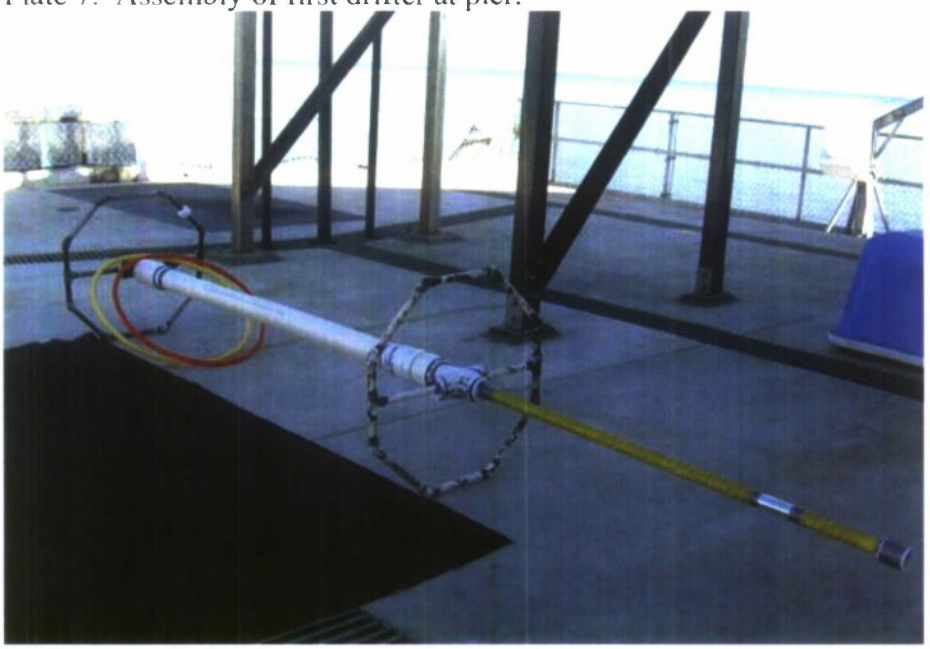
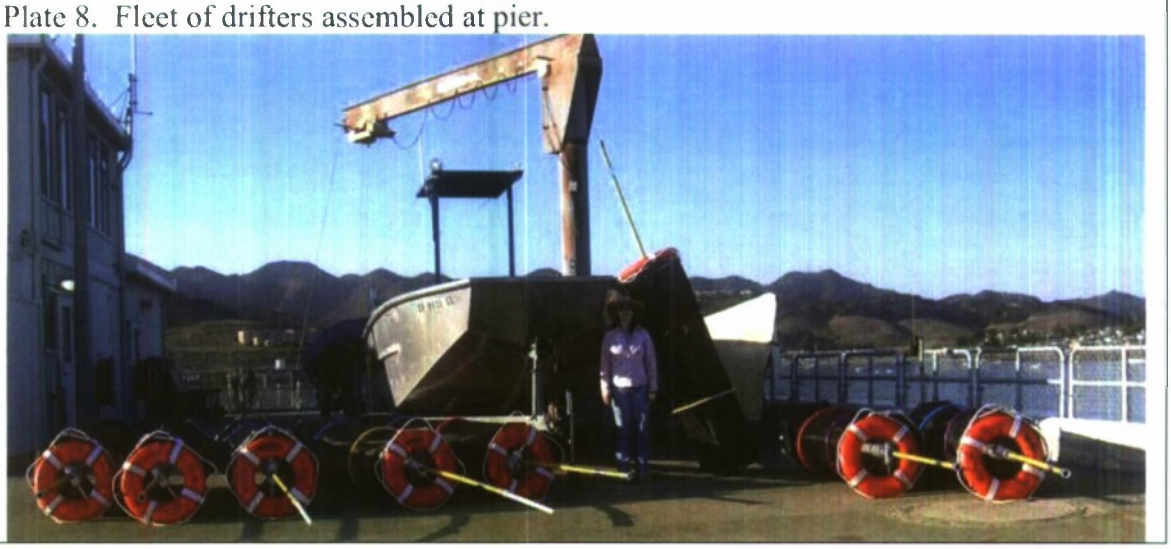


Plate 8. Fleet of drifters assembled at pier.



Seheduling time with the Cal Poly boats and pier manager was ehallenging, since the facility is used by many Cal Poly researehers and others (NASA, ete). We got approval for a run on August 16, and by 9 am placed all eight drifters in the oeean water, close to the Avila Breakwater. The oeean drifters began traveling roughly South-East (see Figure 2 that summarizes the drifter position data gathered during the run). However, soon after the drifters were plaeed in the oeean, the winds pieked up and the waves became so tall that we could no longer safely use Cal Poly's flat-bottom boats. The Harbor Patrol offereд to lend us a hand, so we were able to pick up six out of eight of the drifters

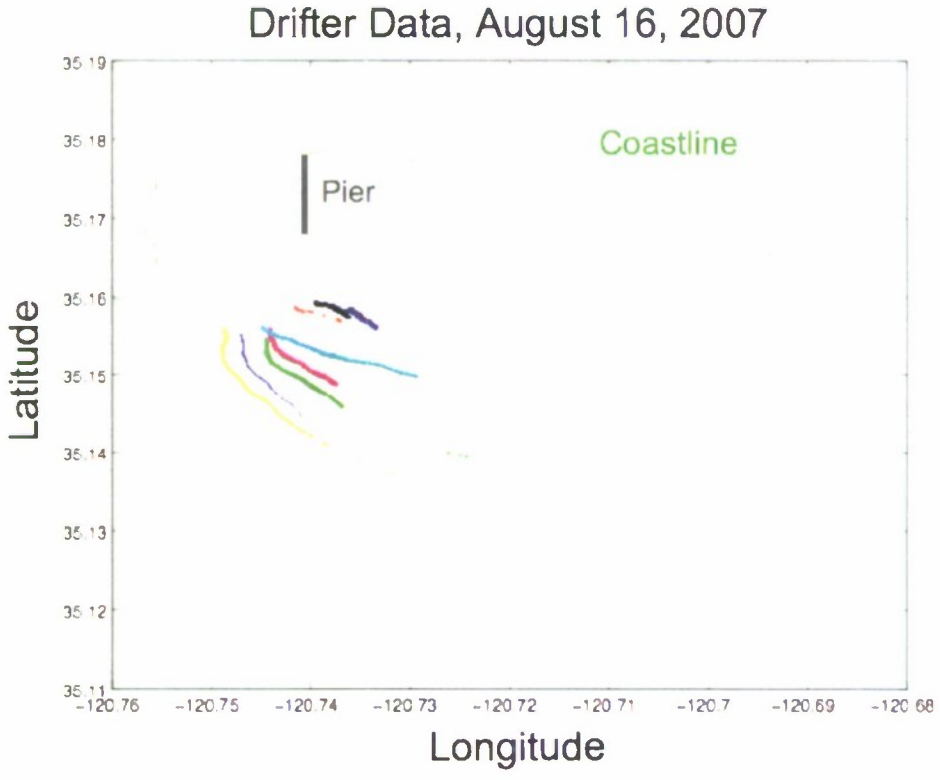


Fig. 2. GPS position data from eight drifters, also called "drifter tracks".

in the gale-forcc winds and extreme wavcs (15 feet tall). A week later, the two remaining "lost" drifters were found and returned to us. (One was found in the Diablo Nuclear Plant exelusion zonc and one near the public Avila pier. Both were floating at the surfacc and had not taken in water.) Thus we observed that the drifters ean withstand galc-forcc winds and high seas. Figure 2 is a summary of our progress in one year of C3RP support. We started from no design or prototype and were able to build a flcet of oeean drifters that were traeked in real time and provided data that eould be plotted in latitude vs. longitude for oeeanographie analysis. The light blue drifter track shows the path of this unit as it floated with the eurrents for about 3 hours. For eomparison in this figure, the Cal Poly pier is one half mile long.

Preliminary Data Analysis.

Figure 3. Wind speed measured at Cal Poly pier from August 12 to 17, 2007.

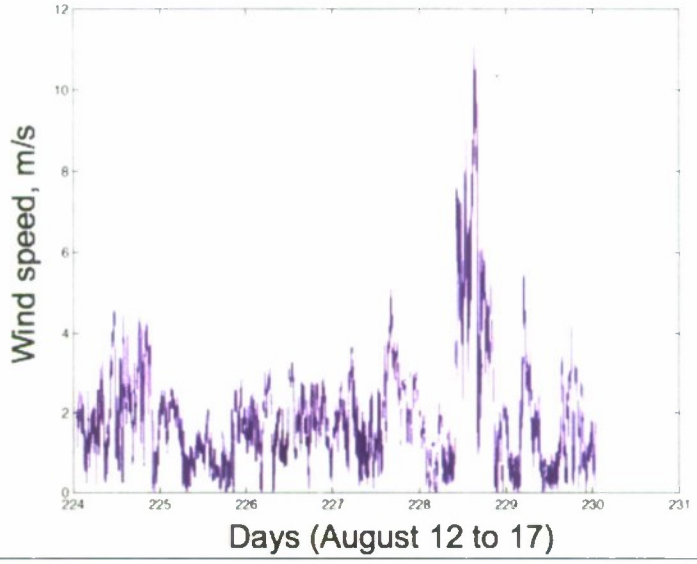


Figure 4. Wind direction measured at Cal Poly pier on the morning of Aug. 16, 2007.

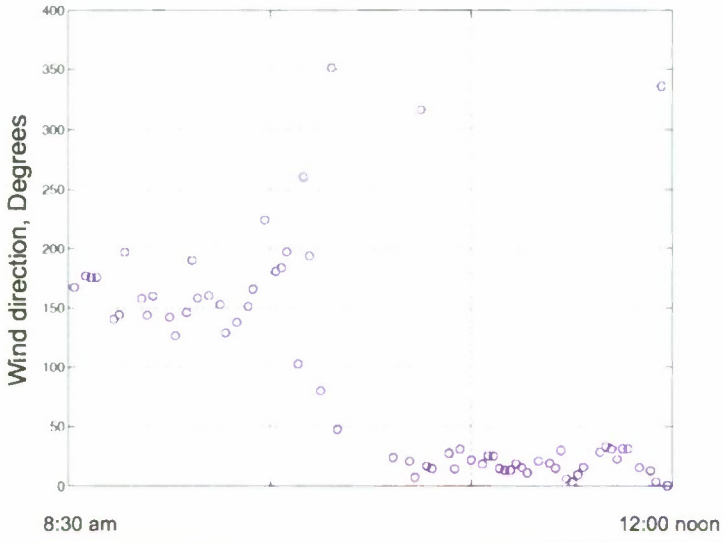


Figure 3 shows the unusually high winds on August 16, 2007. Over the ocean, the winds on August 16 were gale-force and caused such large waves that the Cal Poly's flat-bottom boat could no longer safely transport personnel on the ocean. During that morning, the winds shifted from traveling from the south (180 degrees in Figure 4) to coming from the north (0 degrees). Since the drifter tracks of Figure 2 do not drastically shift direction 45 minutes into data acquisition, we conclude that surface currents on this day were not directly driven by winds. In fact, Figure 5 shows the diurnal (daily) cycling of wind direction over almost a week, as measured at the pier. Our data suggests that the ocean current direction is not cycling in direct response to these diurnal winds.

Figure 5. Wind direction measured at Cal Poly pier over six days.

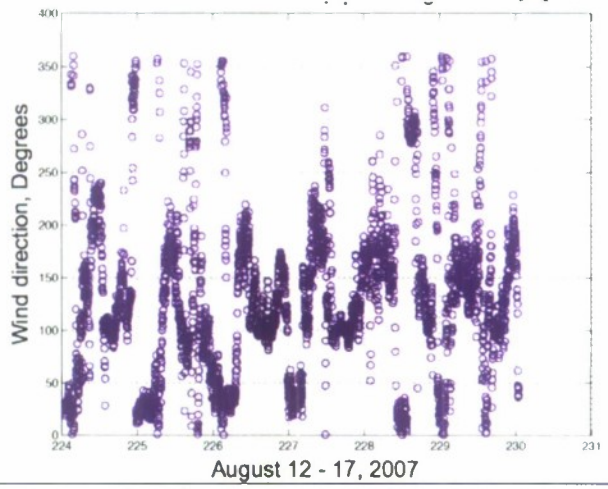


Plate 9. Bathymetry map showing curves of constant depth in San Luis Obispo Bay.

Plate 9 shows the bathymetry (underwater topography) map near Point San Luis, in San Luis Obispo Bay. The curves show locations of equal depth. By comparing Figure 2 and Plate 9, it is apparent that the drifters were traveling approximately along paths of constant depth. Interestingly, on the morning of August 16, 2007, the tide was coming in to shore. So the drifters were not driven out into the ocean by the tides. One possible interpretation is that the drifters may have been beginning to trace a recirculating path within San Luis Obispo Bay. For example, it is known that a prevalent ocean current in this region is the California Current that travels roughly from north to south along the surface waters of the coast. So the drifter paths of Figure 2 may be indicating a recirculating pattern in the bay stemming from the bordering California Current at sea. The main conclusion at this stage is that much more data collection is required to make conclusions from this type of data. In addition, it is unfortunate that high winds forced an early conclusion to the first data collection using the drifters on August 16.

Plan for Future Modifications, Additional Drifters and Data Collection:

The experience taught us several ways to improve our system. In the next funding cycle we plan to: (1) Create a Matlab code that can acquire the data, plot it in real time (overlaid on coastal geography) and perform frequent data backups. This would alleviate navigational issues that slowed us down on the maiden voyage, and prevent data loss. (2) Fix a strain relief issue with the signal cable in the drifters, to prevent intermittent transmissions under high loadings (on rough seas).

We will also create an alternate type of drifter that would not need to be picked up at the end of the day, namely a data-logging type. This style of drifter could be easily deployed in San Luis Obispo Bay, and left to drift for long periods, storing the data inside the drifter. We plan to attach salinity and temperature sensors to this style of drifter to obtain data on possible mechanisms driving the flows.

Our data collection strategy is to deploy the original fleet of drifters (8 units) but incorporating a longer battery life so that we can gather much larger data sets within about 10 miles of the pier. We will deploy these over multiple days on many multiple runs, examining the variation over weeks and seasons (for example summer versus spring currents). The second fleet (8 drifters) will also be deployed, and allowed to drift. These will float with the surface currents and potentially map the ocean current behavior over much larger regions (perhaps past Santa Barbara or north to Cayucos for instance.)
In short, we plan to thoroughly utilize both drifter fleets to characterize the surface currents in and around San Luis Obispo Bay.

Data analysis:

Following the analysis approaches of modern oceanographers that use drifters, we will create maps of the time-average surface currents, and maps of mean and eddy kinetic energy along the coast. This will help us to understand the evolution of the flow including its turbulent structure. Finally, an important transport statistic that we will determine from the data is the water residence time in the bay and at other positions along the coast. This parameter is important so that biologists can quanify the living environment of organisms living in the ocean such as phytoplankton and fish. By analyzing the flow with respect to various parameters (wind, tides, bathymetery, etc) we plan to improve modeling efforts of coastal flows. This can also contribute to creating better estimates of the required time-in-the-water for military personnel swimming to foreign beaches.

Experimental and Numerical Analysis: Tactile RPG Active Protection System

Project Investigators:

Garrett J. Hall and Eric P. Kasper
Civil and Environmental Engineering
California Polytechnic State University
San Luis Obispo, CA

Section 1

Introduction

The intent of this report is to convey the results of an experimental and numerical analysis conducted on the Tactile RPG Active Protection System (TRAPS). The simulations focused on the determination of the response of the radar box with and without isolation, see Figure 1.1. The primary response quantities of interest are the fundamental frequencies and mode shapes, the experimental response under a PSD base motion, and the effect of vibration isolation of the radar box. An outline of the report is as follows.

Section 2 describes the modeling approach taken in the experimental phase of the system and consists of the following tasks:

- Develop a preliminary finite element model: Develop a finite element model of the physical system (i.e. TRAPS) with the intent of utilizing the numerical model to help guide the experimental phase. Specifically, numerically determining the mounting fixture geometry for the shake table such that the interaction between the TRAPS and the mounting fixture is minimized. This step will aid in obtaining an accurate representation of the natural frequencies and time history by accounting for uncertainties associated with the mounting fixture.

- Build the test mounting fixture: Given the results from the preliminary finite element analysis fabricate the mounting fixture that will allow for motions to be imposed on the TRAPS for all three axes.

- Experimental sine sweep simulation: Perform an experimental

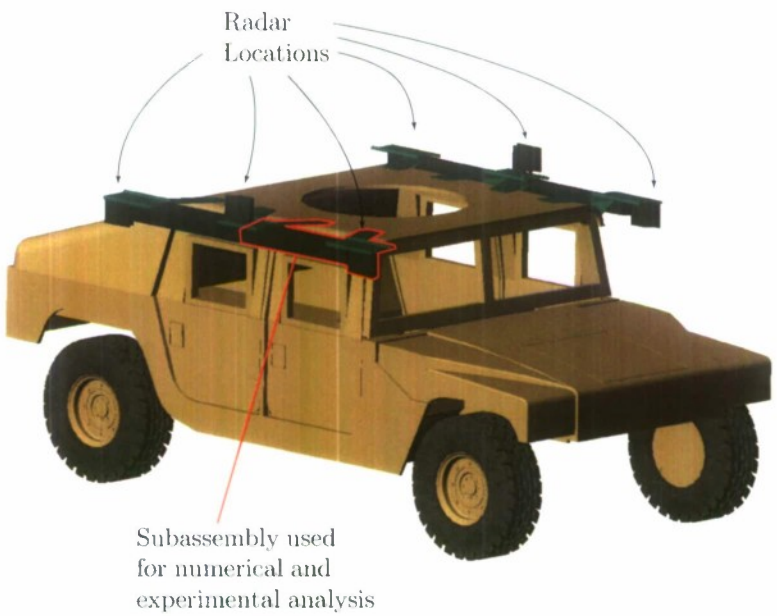


Figure 1.1: Overview of the vehicle showing TRAPS radar locations and subassembly.

sine sweep in three orthogonal directions to determine the fundamental frequencies and also the damping characteristics of the system. Perform the simulation with and without the isolators in place.

- **Experimental PSD simulation:** Using the provided MIL-STD-810E minimum integrity test as input for the shake table run the simulation in three orthogonal directions to determine the accelerations of the radar box with and without the isolators in place. In addition include considerations for video taping of the tests.
- **Experimental testing of isolators:** Perform a suite of uniaxial tension/compression experimental tests for the isolator components. The uniaxial test will be performed at three different loading rates to assess the effect rate has on the response.

Section 3 describes the modeling approach taken in the numerical analysis of the system. The numerical phase of the project consists of the following tasks:

- **Numerical model for the isolator:** Given the experimental data determined from the suite of uniaxial tension/compression tests develop a simple non-linear model (*i.e.* connector element) which captures the macro behavior of the isolators.
- **Refined finite element model:** Utilizing the preliminary model developed for the experimental phase and the connector model perform convergence studies for both spatial discretization (mesh density) and temporal discretization (time step) to minimize the uncertainty associated with the numerics. Particular attention is paid to understanding the sensitivity of the analysis to the modeling assumptions so that one has a realistic assessment of the potential variability in the analytical results. Subsections detail the theoretical assumptions, the modeling assumptions, mesh density, modal dynamic inclusion, connection models, and boundary conditions.
- **Numerical sine sweep simulation:** Utilizing the refined finite element model and the data from the experimental tests (*i.e.* determination of damping characteristics) perform a set of finite element simulation of the sine sweep in three orthogonal directions and compare with key acceleration data at various points on the structure.

Perform the simulations with and without the isolators to determine their effect on the response.

The final section of the report summarizes the findings of the study. By comparing the experimental work with the numerical model, in the context of the modeling sensitivities outlined in Section 2, one can assess the uncertainty in the results and judge the ability of the numerical approach to be used as a guide for future design iterations.

Section 2

Experimental Phase

2.1 Introduction

In this section the experimental protocol for the analysis of the TRAPS radar box is presented. To enable an accurate representation of the boundary conditions of the TRAPS a preliminary finite element model was developed. The numerical model will give an indication of the influence the support structure has on the response. Ideally the support structure should not impact the response, but would mimic a rigid connection in order to identify the TRAPS response. Once the geometric specification of the support structure is completed the next stage is to fabricate the test setup. The fabricated support structure is then mounted on the shake table along with the TRAP system.

The first step is to perform sine sweeps in three orthogonal directions to assess damping in the system and determine the fundamental frequencies. Next a base input motion, developed from information provided by Textron, will be used as input for the shake table. The base motion will be performed in three orthogonal directions. Acceleration and video data was collected for all runs for subsequent comparison with the numerical models. For each experimental simulation the radar box will be examined with and without rubber isolators to assess the performance of the isolators.

2.2 Preliminary Finite Element Model

A finite element model was developed with the intent of utilizing the numerical model to help guide the experimental phase. Specifically, the model will assist in numerically determining the mounting fixture geometry for the shake table such that the interaction between the TRAPS and the mounting fixture is minimized. By obtaining the correct boundary conditions (*i.e.* support structure) an accurate representation of the natural frequencies and time history may be determined along with uncertainties associated with the mounting fixture.

Due to the topology of the parts (*i.e.* two predominate dimensions relative to the thickness direction) shell elements were utilized for all parts. For the given geometry shell elements will accurately capture both the in-plane and out-of-plane (bending) response. The shells elements used were linear four-node quadrilateral elements using four point gauss quadrature in the plane of the element and five point simpson's rule through the thickness of the element. The preliminary model contains 97500 degrees-of-freedom. An overview of the finite element model is shown in Figures 2.1-2.2 .

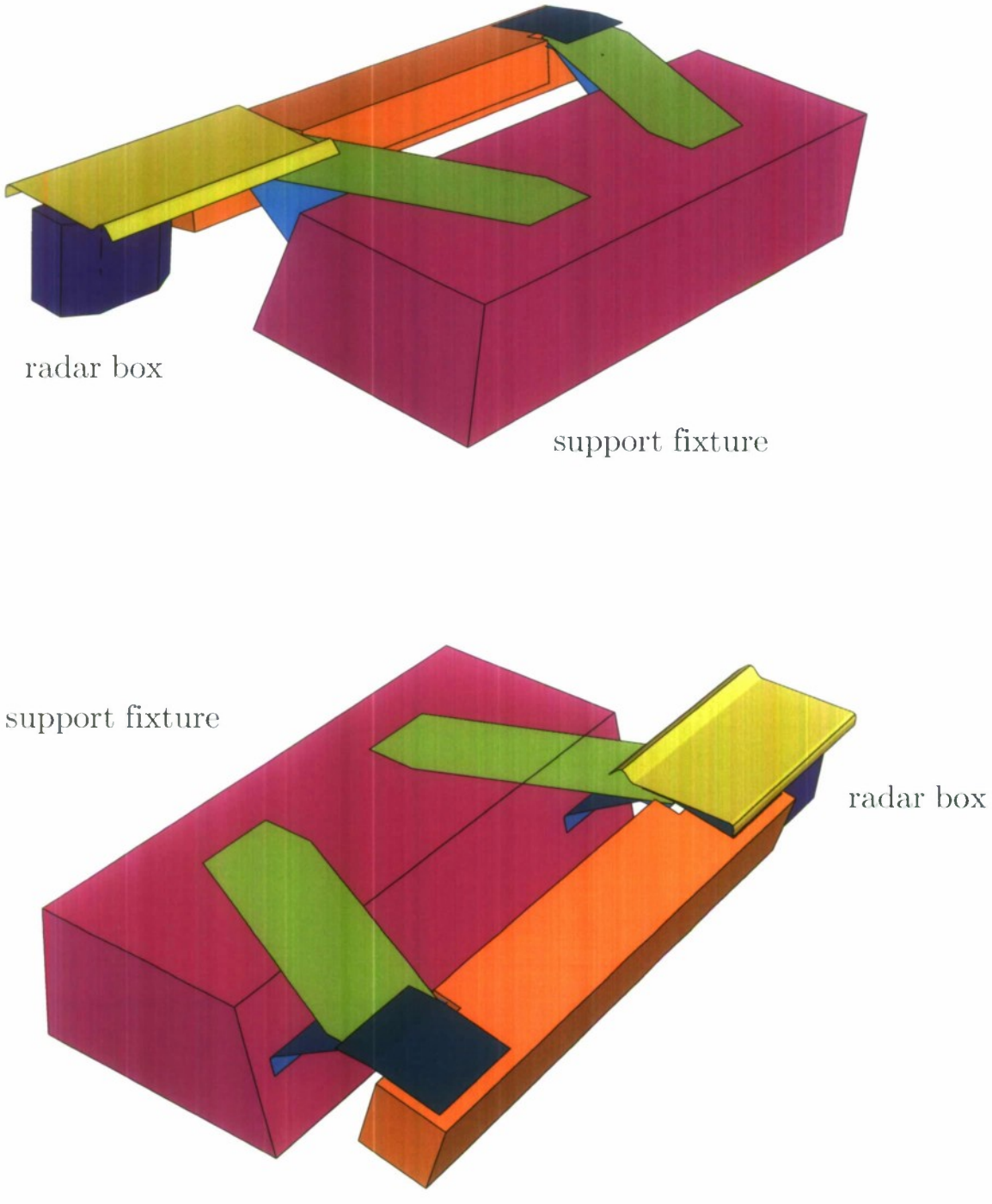


Figure 2.1: Perspective views of the model with parts color coded.

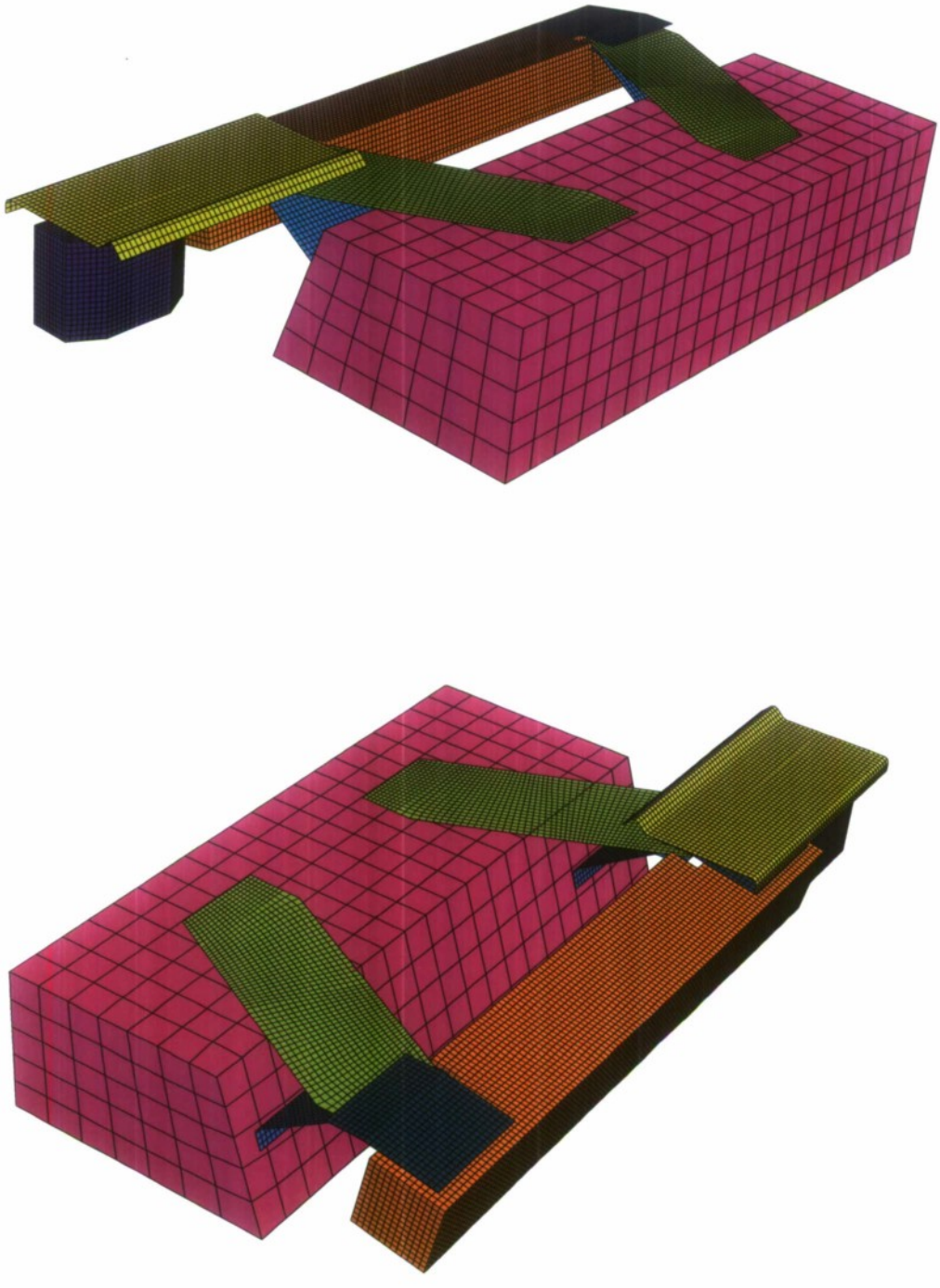


Figure 2.2: Perspective view of the model showing a sample mesh.

2.2.1 Support Fixture Analysis

To determine the geometric specifications for the support fixture an eigenvalue analysis was performed on the counter measure box and radar system in which the boundary conditions were assumed perfectly fixed, see Figure 2.3. A second eigenvalue analysis was then performed which included the support fixture. The support fixture is fixed at its base at several locations to mimic bolting the support to the shake table, see Figure 2.4. Again the goal of these simulations was to minimize the effect the support fixture has on the response of the counter measure and radar box.

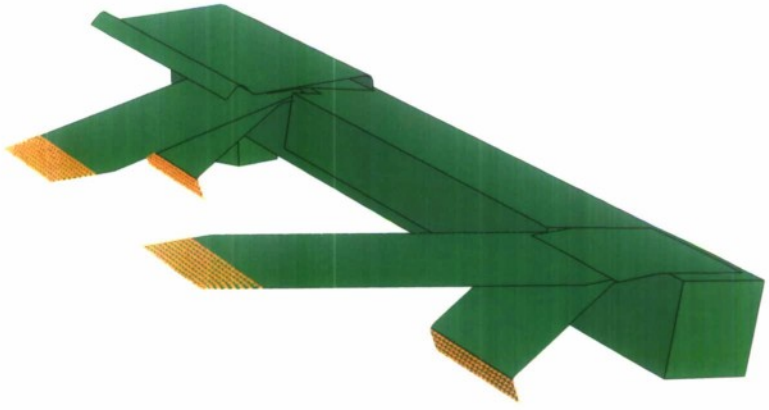


Figure 2.3: Counter measure and radar box without support fixture: boundary conditions highlighted at the ends of the roof bracket supports.

As summary of the first ten natural frequencies for both configurations is shown in Table 2.2.1 and accompanying modes shapes are shown in Figures 2.5 - 2.14. The frequencies and the mode shapes below indicate that the designed box will adequately represent a rigid attachment point to the shake table.

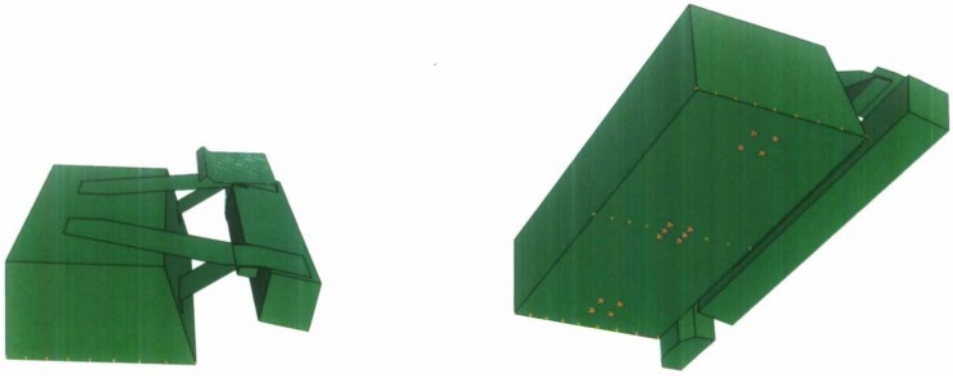


Figure 2.4: Counter measure and radar box with support fixture: boundary conditions highlighted at the bottom of the support fixture.

Table 2.1: Summary of first ten frequencies for structure with and without support box.

mode	fixed support (Hz)	box support (Hz)	error
1	14.090	14.075	0.11 %
2	25.500	25.483	0.07 %
3	35.819	35.653	0.46 %
4	47.607	47.418	0.40 %
5	63.410	63.399	0.02 %
6	141.88	139.52	1.66 %
7	159.94	157.29	1.66 %
8	164.18	162.11	1.26 %
9	174.67	174.15	0.30 %
10	196.63	196.09	0.27 %

2.3 Fabrication of Mounting Fixture

Based on the results from the preliminary finite element analysis a mounting fixture box was constructed. The mounting fixture was installed on the shake table shown in Figure 2.15.

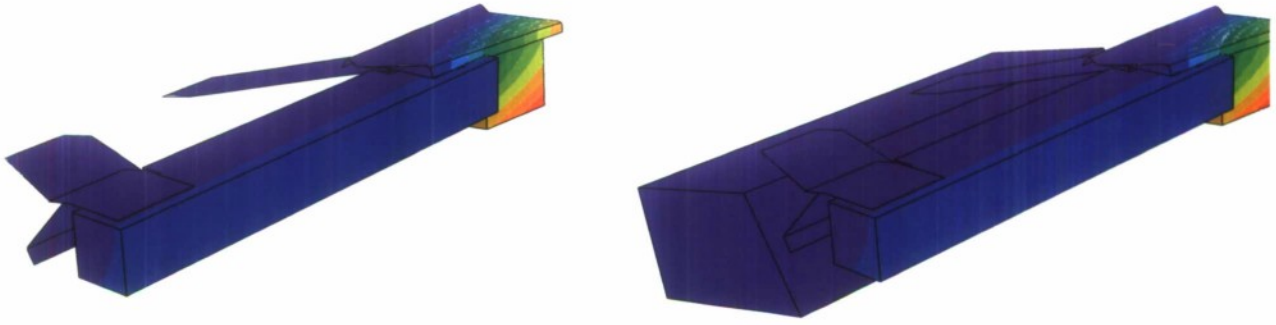


Figure 2.5: Comparison of first mode with and without support fixture.

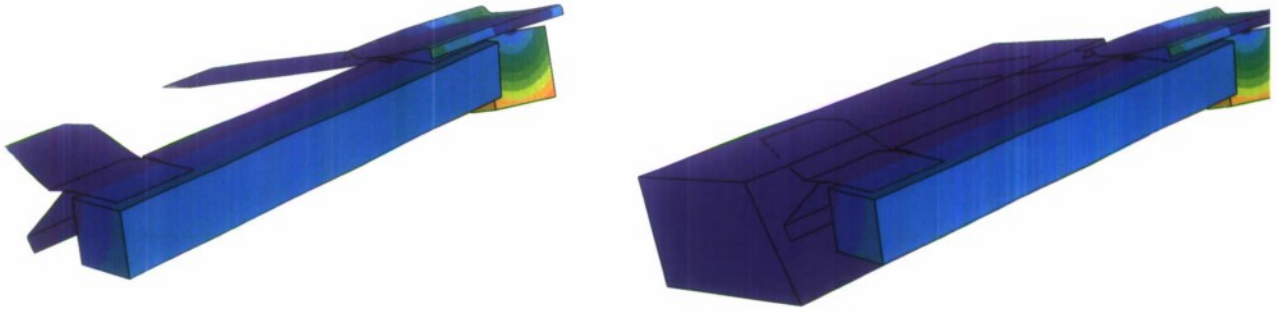


Figure 2.6: Comparison of second mode with and without support fixture.

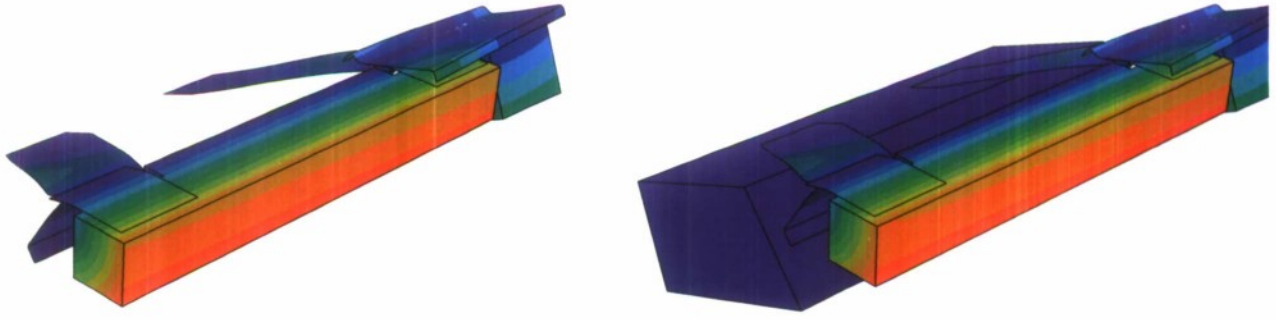


Figure 2.7: Comparison of third mode with and without support fixture.

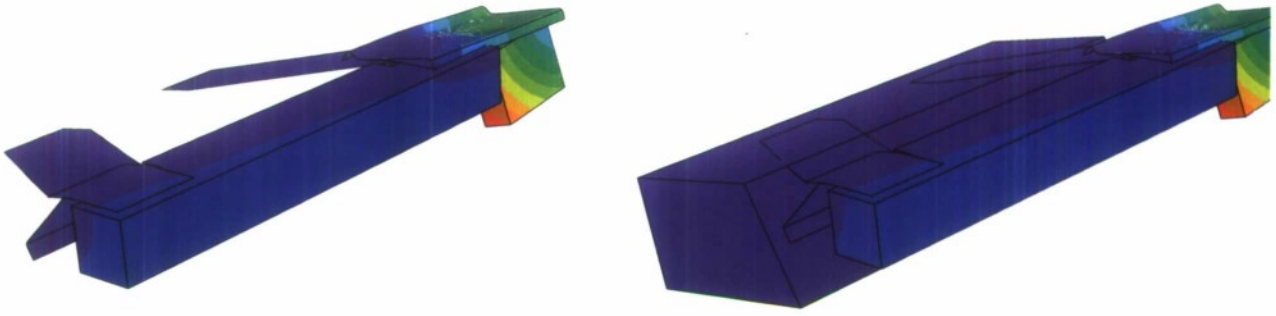


Figure 2.8: Comparison of fourth mode with and without support fixture.

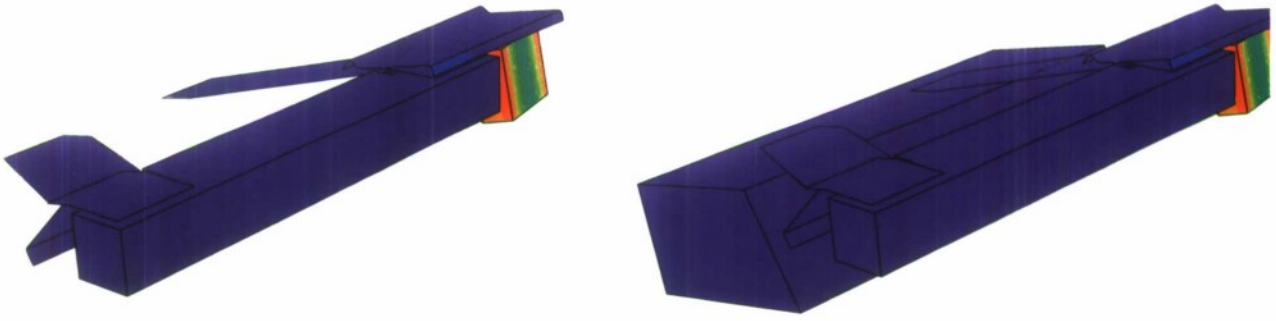


Figure 2.9: Comparison of fifth mode with and without support fixture.

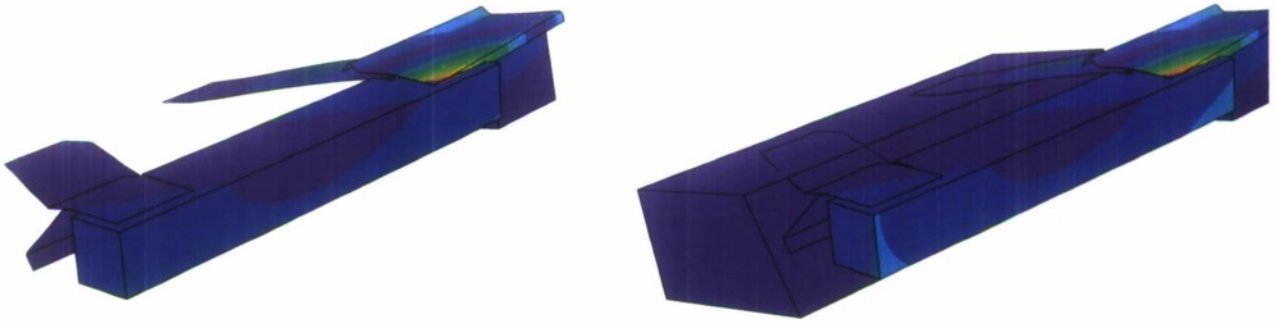


Figure 2.10: Comparison of sixth mode with and without support fixture.

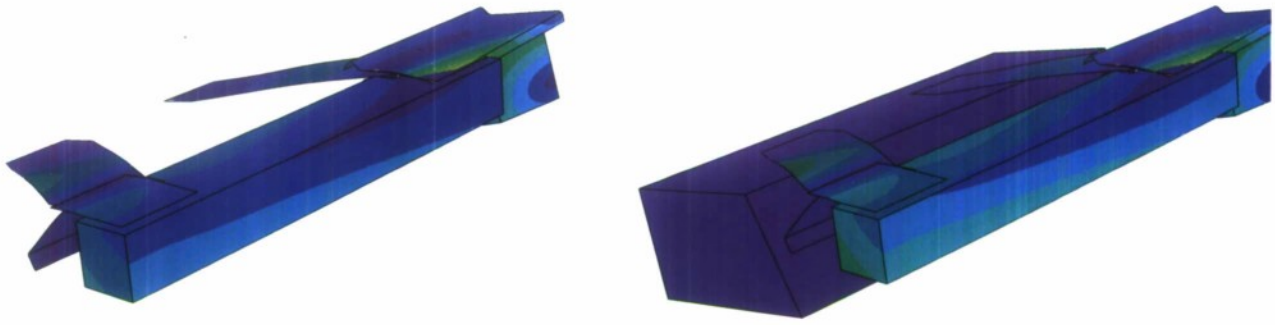


Figure 2.11: Comparison of seventh mode with and without support fixture.

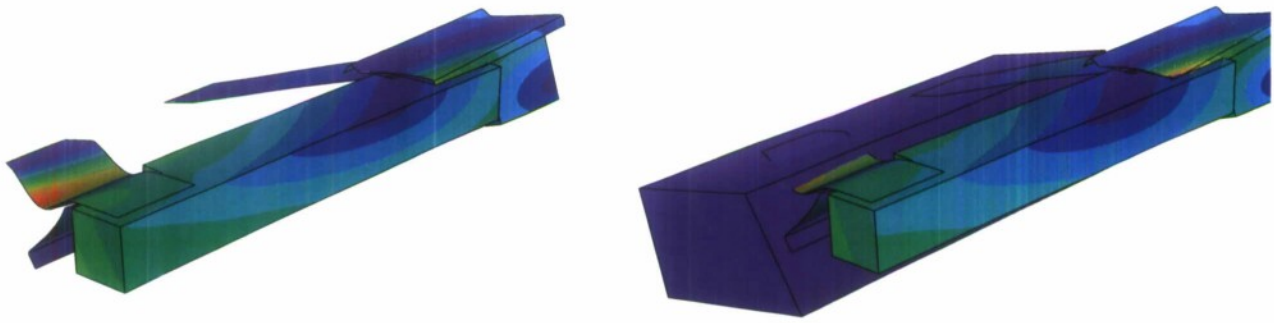


Figure 2.12: Comparison of eight mode with and without support fixture.

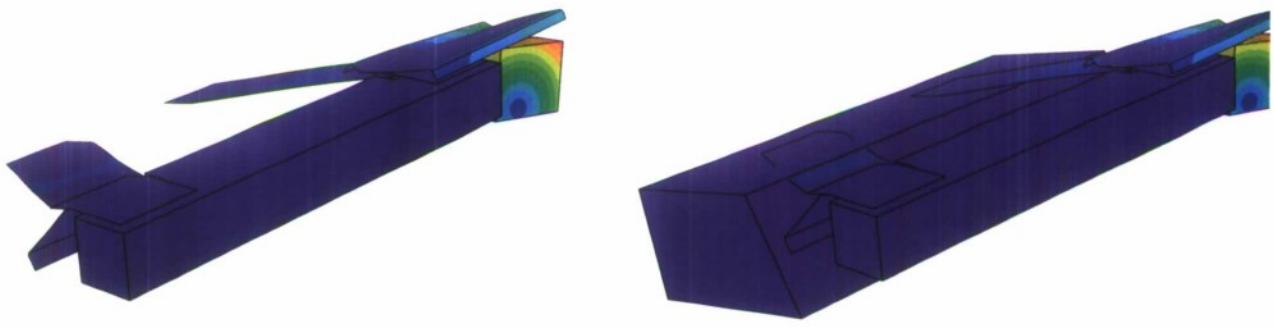


Figure 2.13: Comparison of ninth mode with and without support fixture.

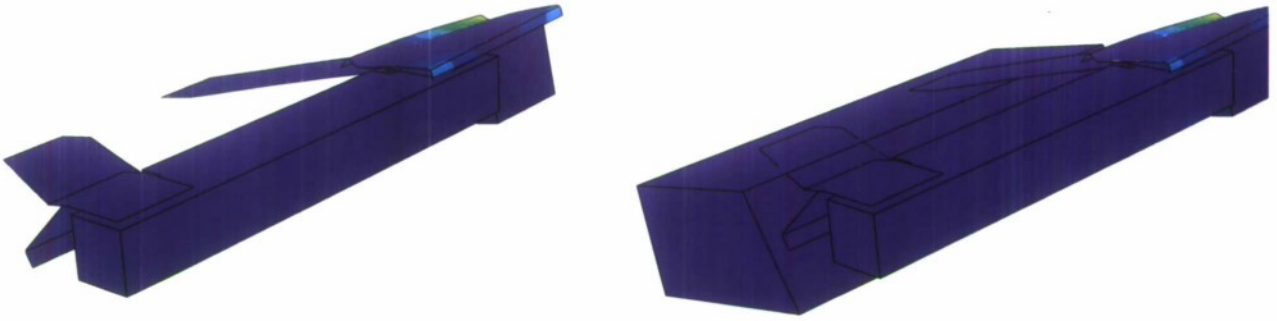


Figure 2.14: Comparison of tenth mode with and without support fixture.

Figure 2.15: Mounting box fixture.

Figure 2.16: Mounting box fixture with counter measure.

2.4 Experimental Sine Sweep Simulations

The mounting fixture box is attached to the shake table shown in Figure 2.17. The table will then programmed to a perform sine sweep, with the test repeated along three orthogonal directions with respect to the TRAPS. The information collected from these simulations will provided fundamental frequencies as well as modal damping information.

Figure 2.17: Single axis shake table.

An overview of the three orthogonal directions in which the TRAPS was orientated, along with accelerometers is shown in Figures 2.18-2.20. A summary table of the first three natural frequencies are given in Table 2.2. The corresponding response curves under the sine sweep input motion follow in Figures 2.21-2.26.

mode	experimental (Hz)	numerical (Hz)	error
1	15.2	14.1	7.2%
2	28.6	25.5	10.8%
3	36.3	35.7	1.7%

Table 2.2: Summary of natural frequencies without the isolator.

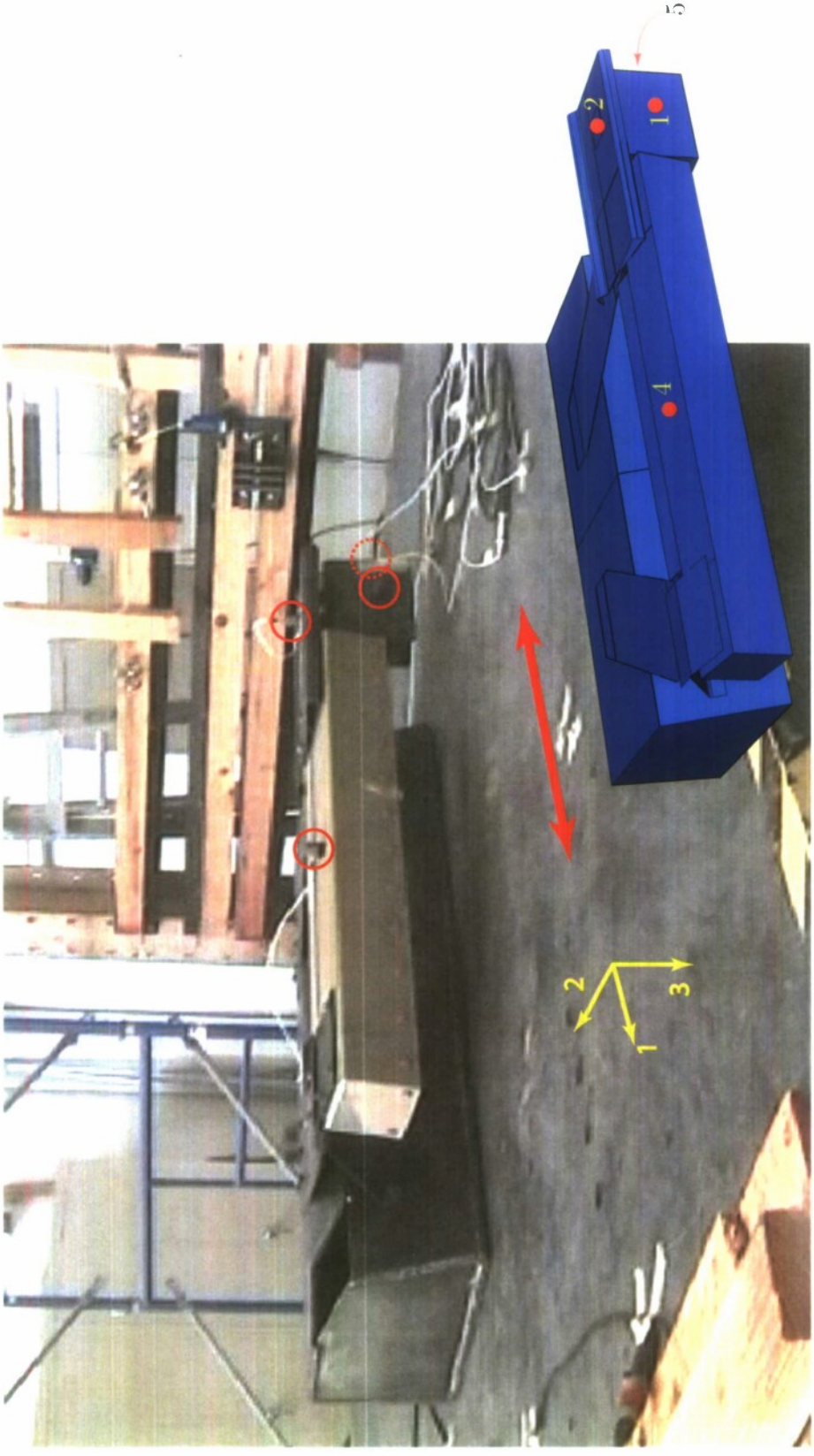


Figure 2.18: Orientation for shaking in the x_1 coordinate axis.

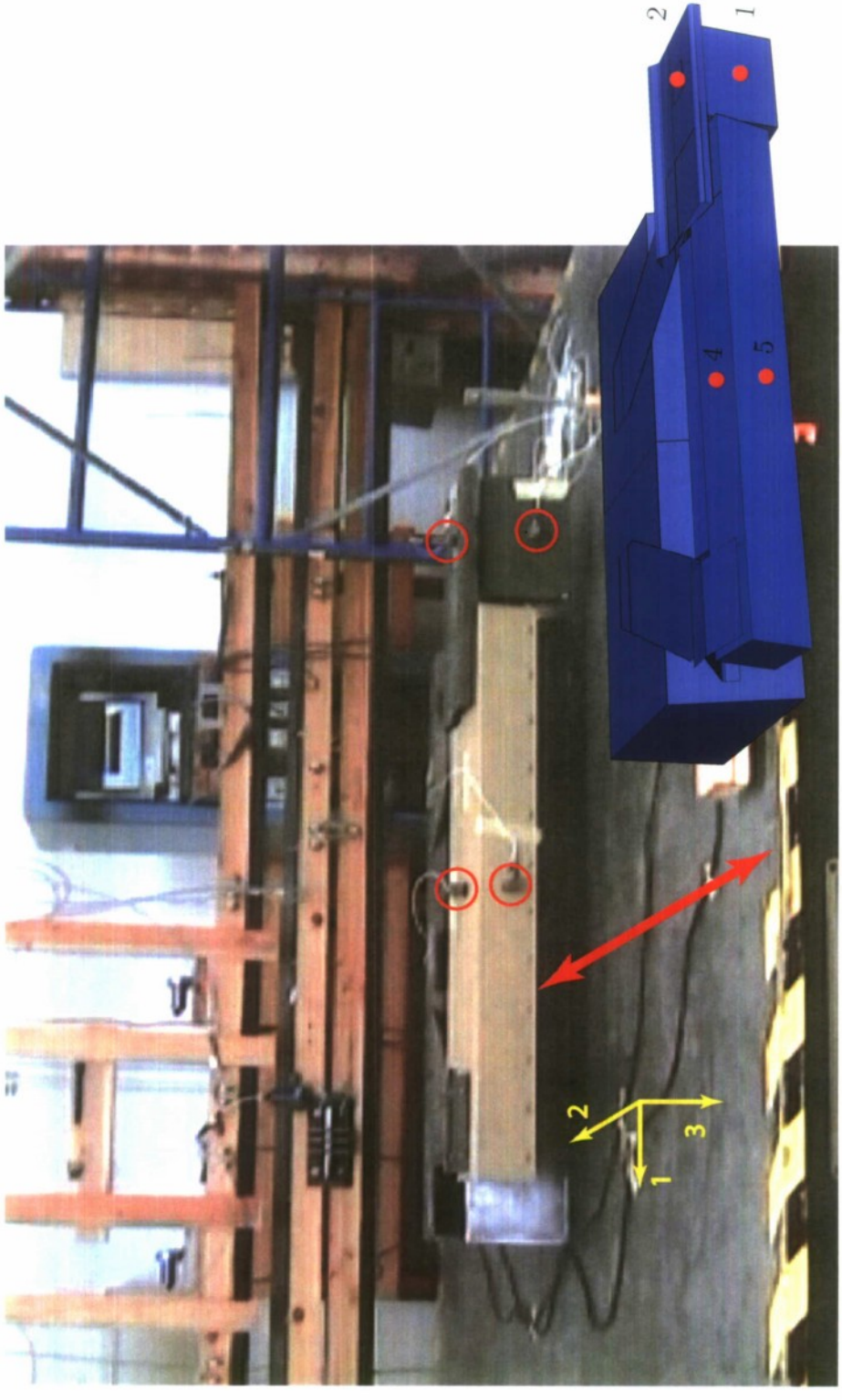


Figure 2.19: Orientation for shaking in the x_2 coordinate axis.

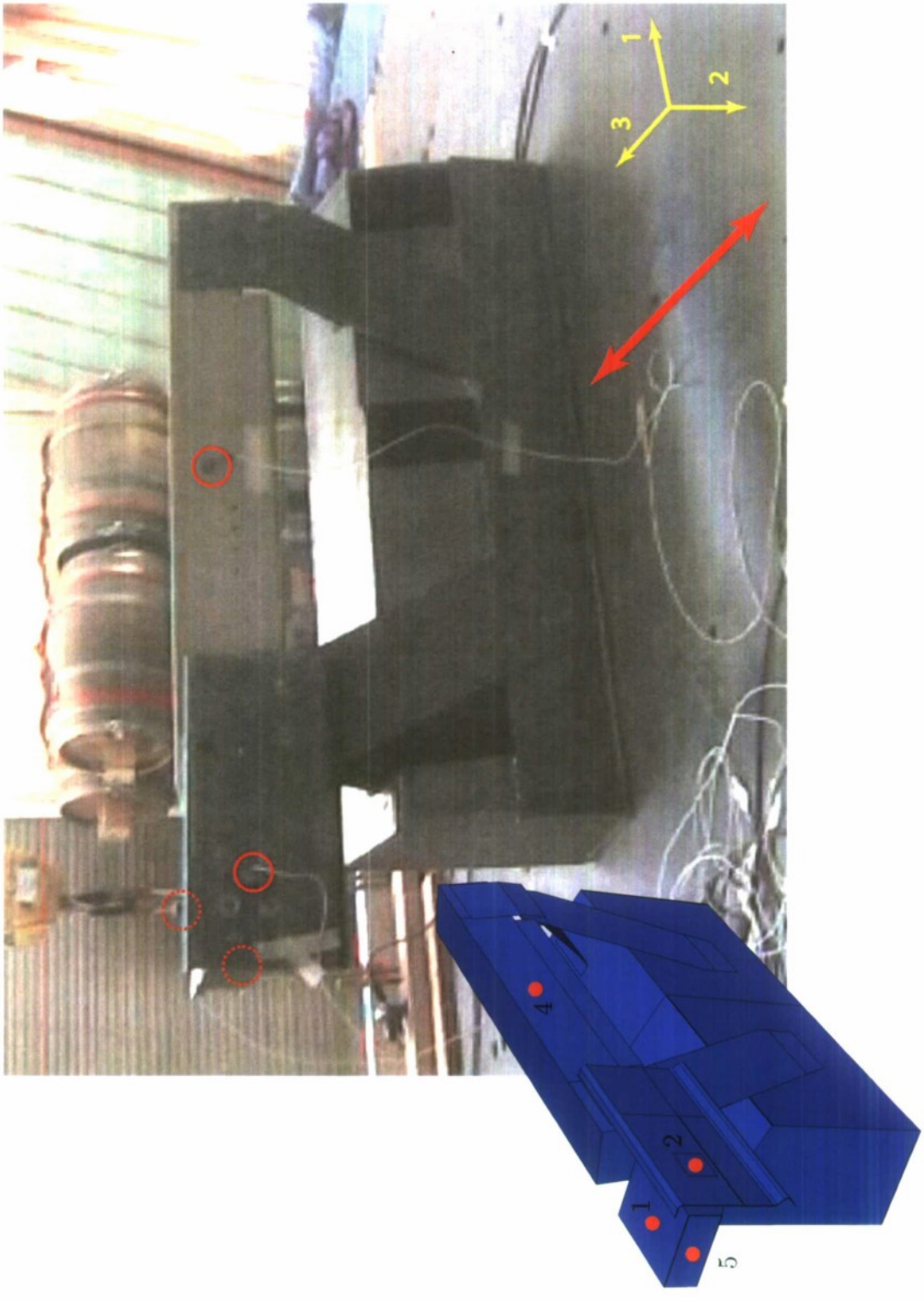


Figure 2.20: Orientation for shaking in the x_3 coordinate axis.

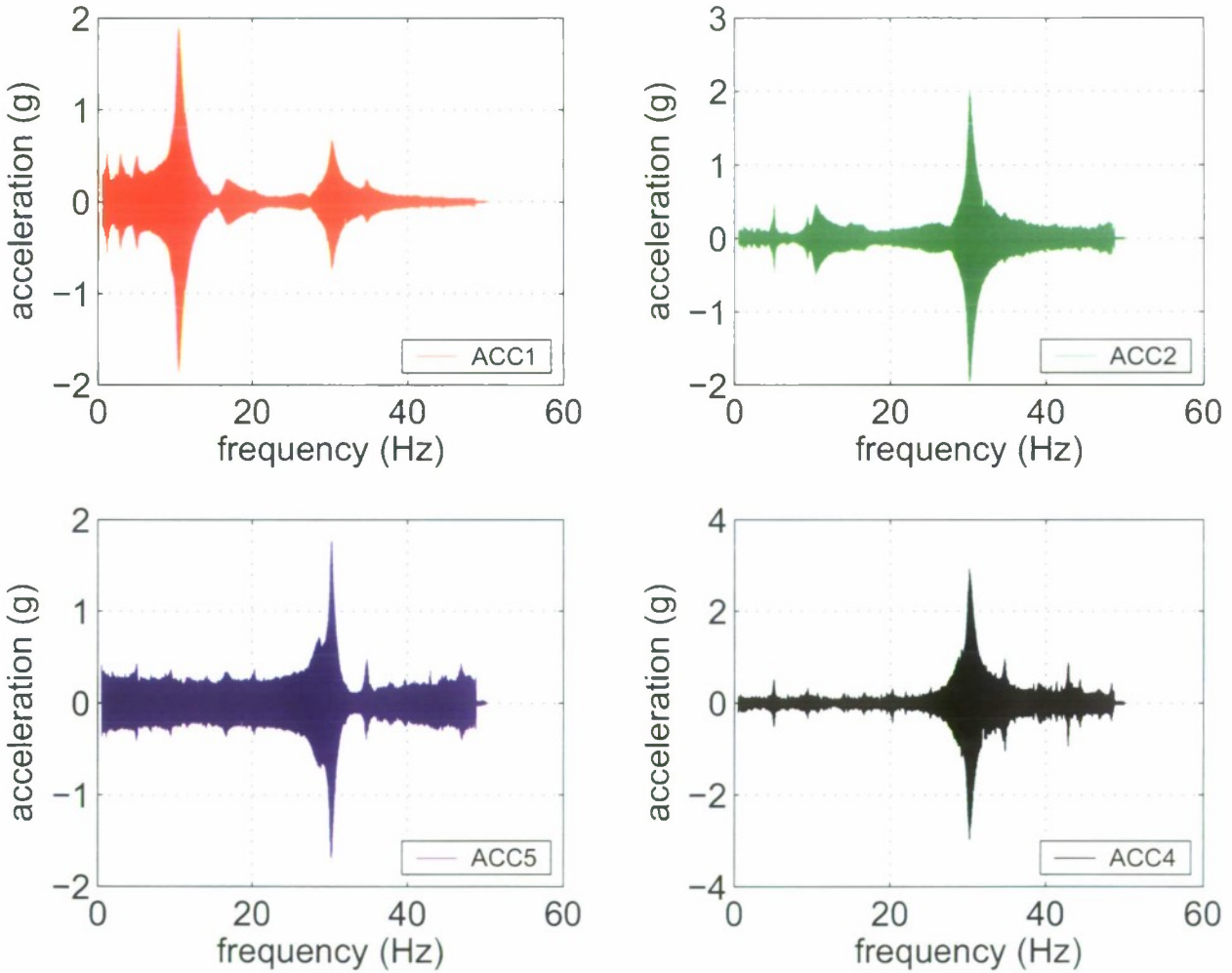


Figure 2.21: Response to sine sweep along x_1, with isolator.

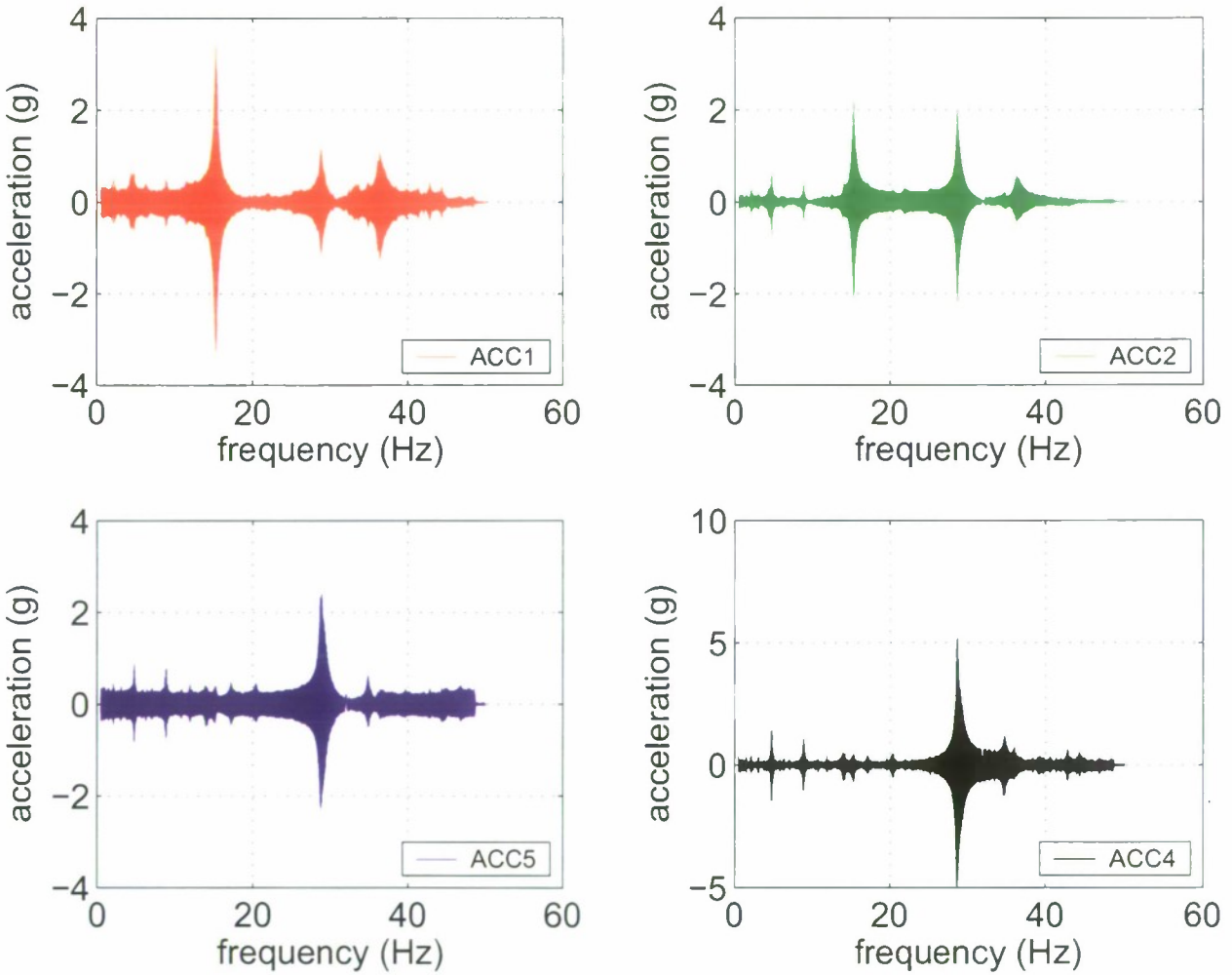


Figure 2.22: Response to sine sweep along x_1, no isolator.

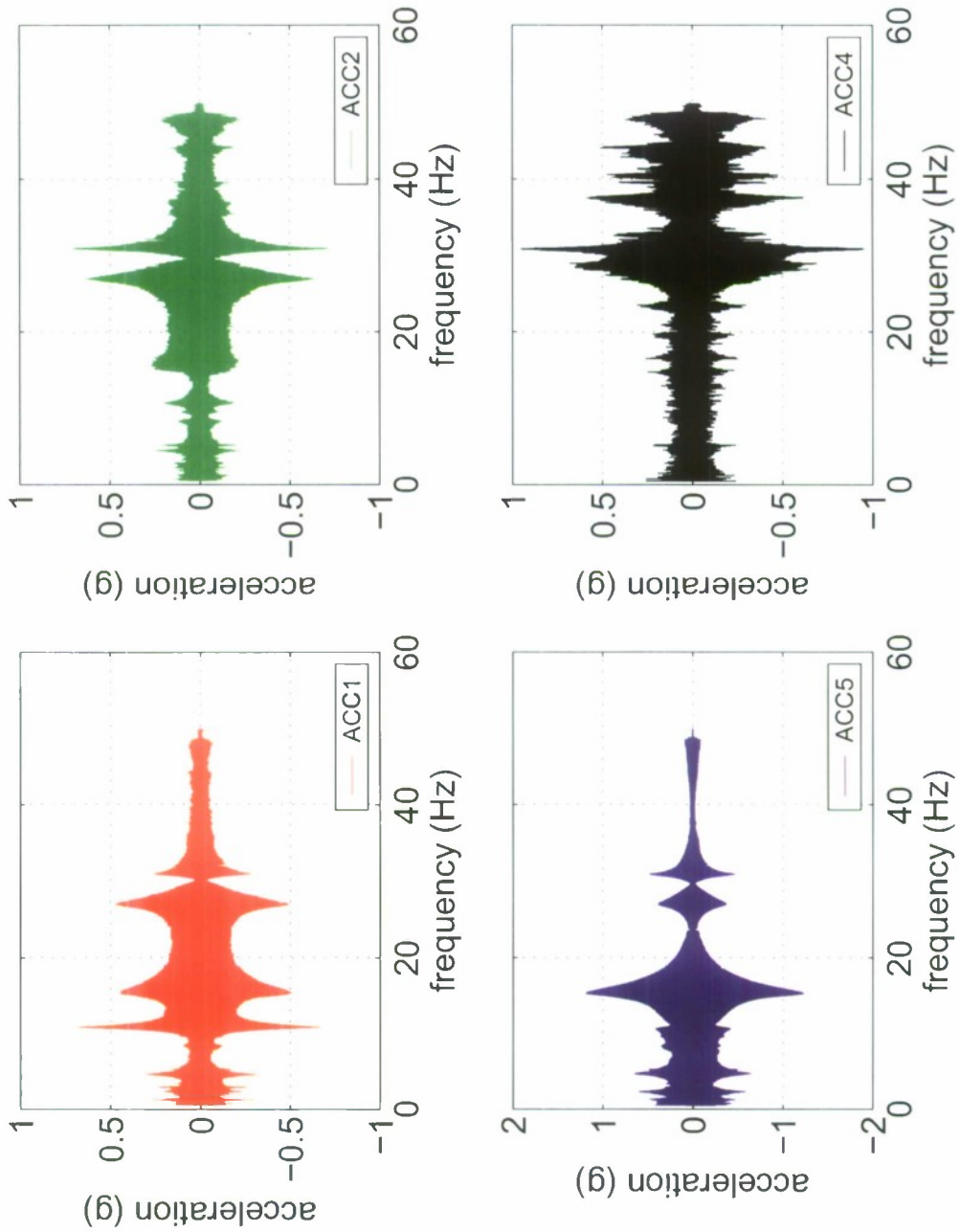


Figure 2.23: Response to sine sweep along x_2, with isolator.

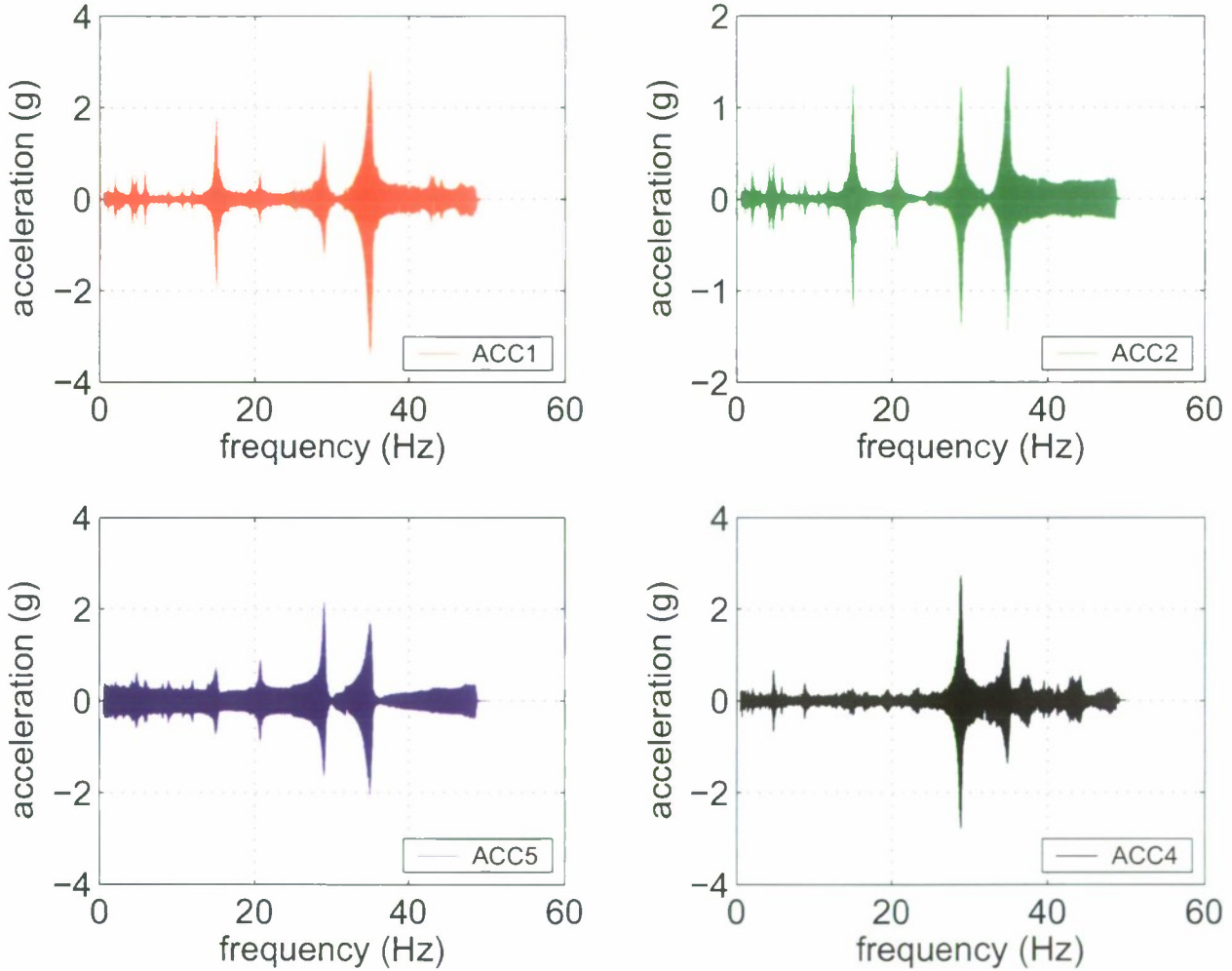


Figure 2.24: Response to sine sweep along x_2, no isolator.

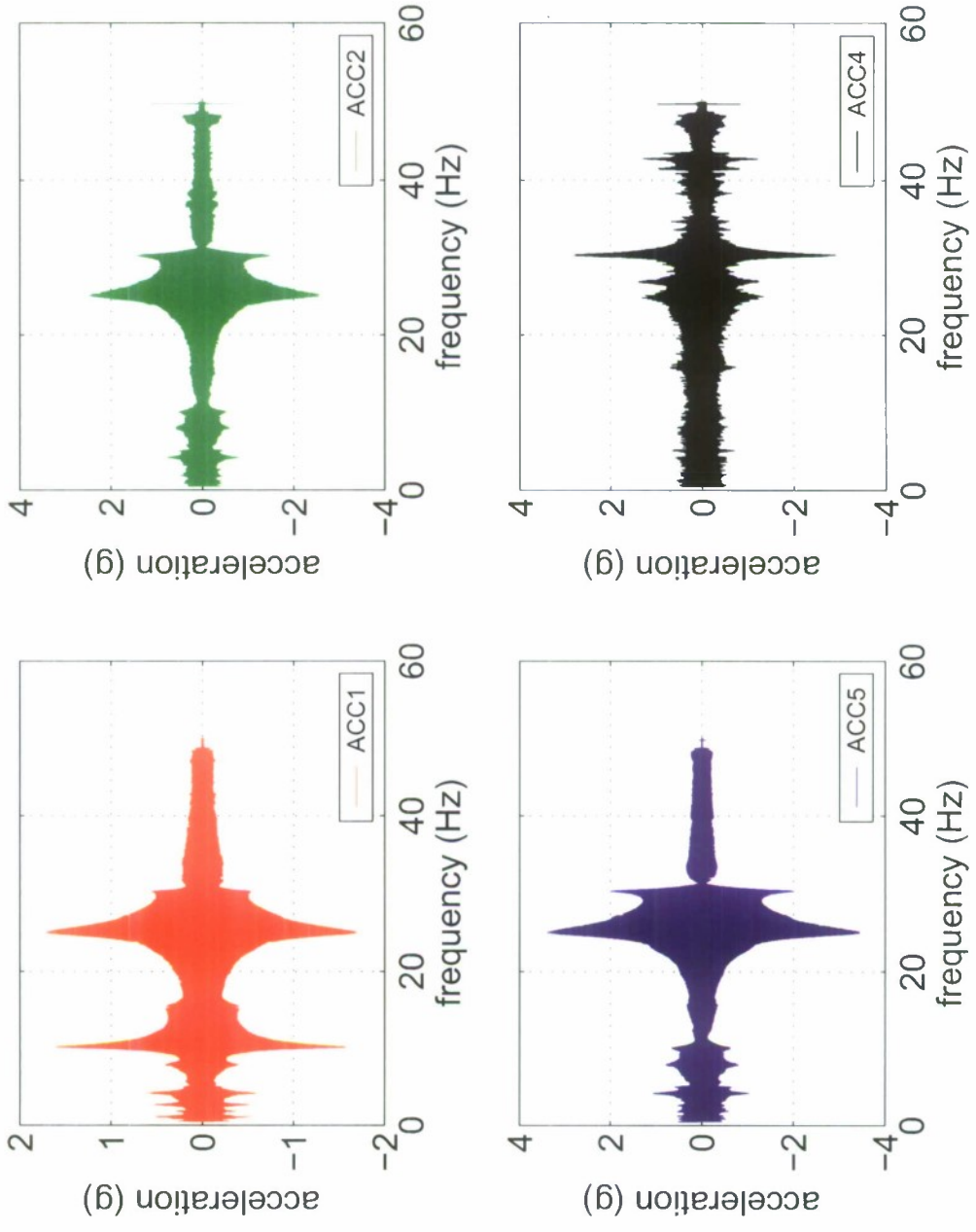


Figure 2.25: Response to sine sweep along x_3, with isolator.

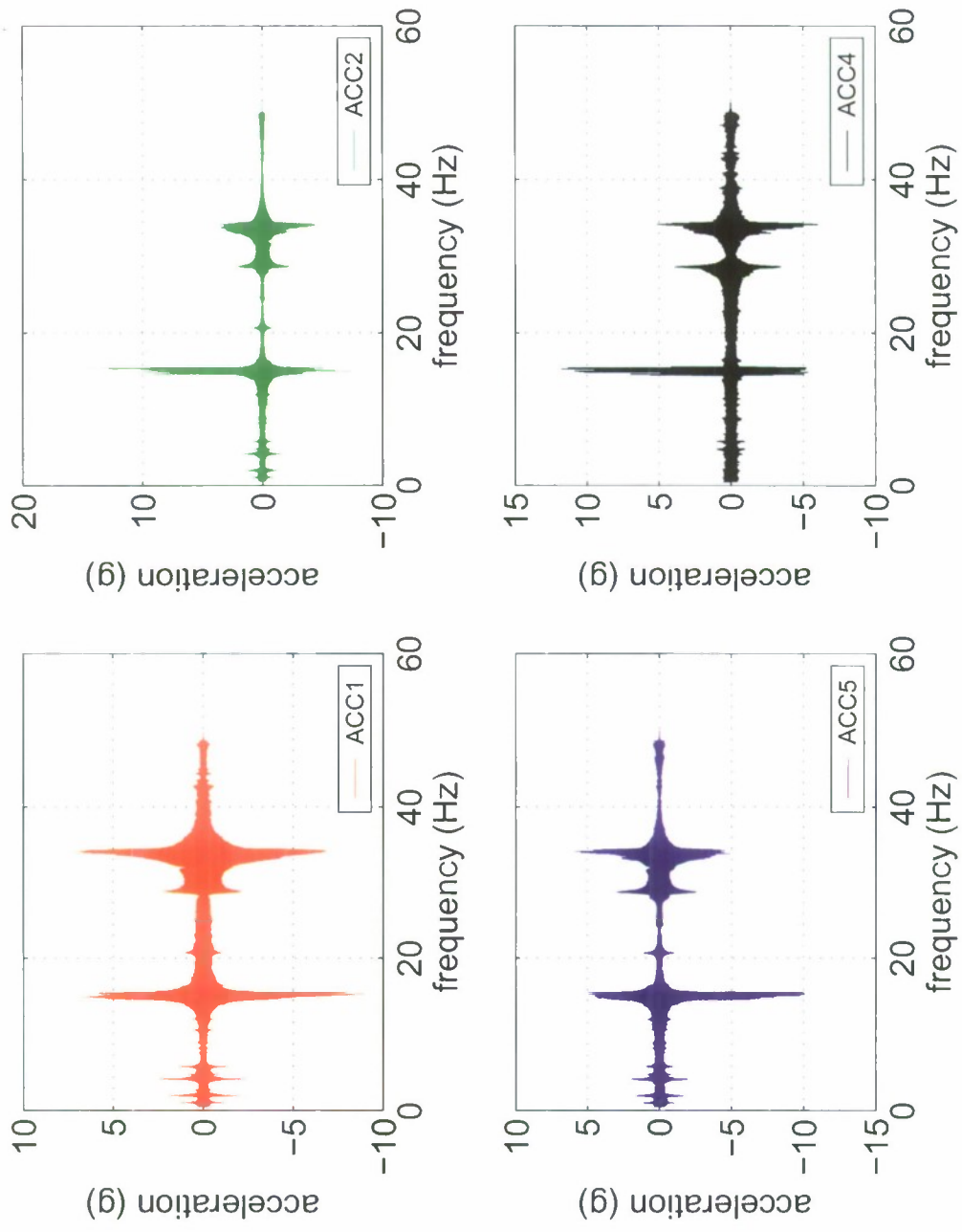


Figure 2.26: Response to sine sweep along x_3, no isolator.

2.5 Experimental Support Motion Simulations

Following the sine sweep experiments, a full support motion was applied to the TRAP system both with and without radar isolator mounts. The support motion will be based on military specifications provided by Textron. Acceleration data was collected in orthogonal directions on the radar box to assess the quantitative effect of the isolators. Video data was also collected to assess qualitatively the performance.

Figure 2.27 gives an overview of the vertical axis shake table with the TRAPS system mounted, while Figure 2.28 depicts the MIL-STD-810E minimum integrity test that is the imposed ground motion. The results from the vertical shake table for the TRAPS with and without the isolators in place are given in Figure 2.29.

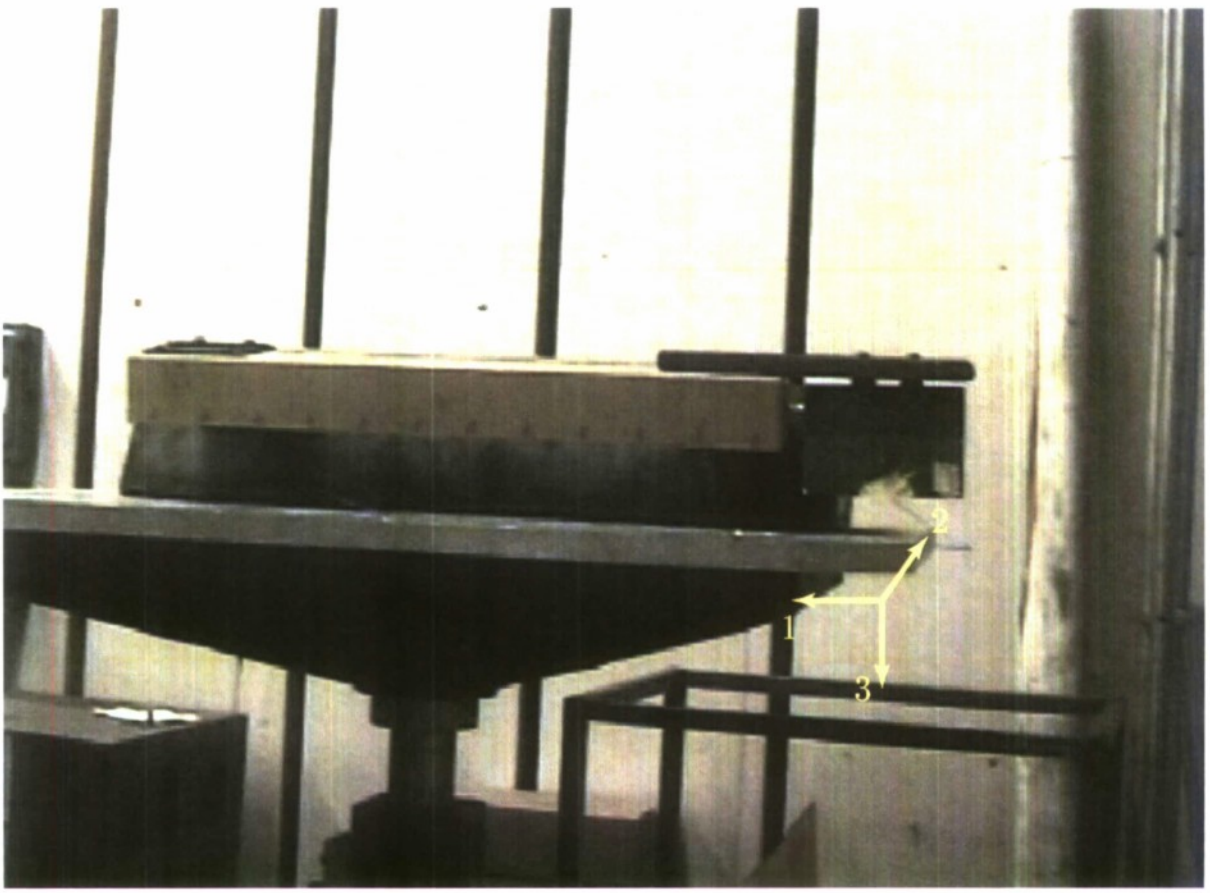


Figure 2.27: Picture of the vertical axis shake table and TRAPS.

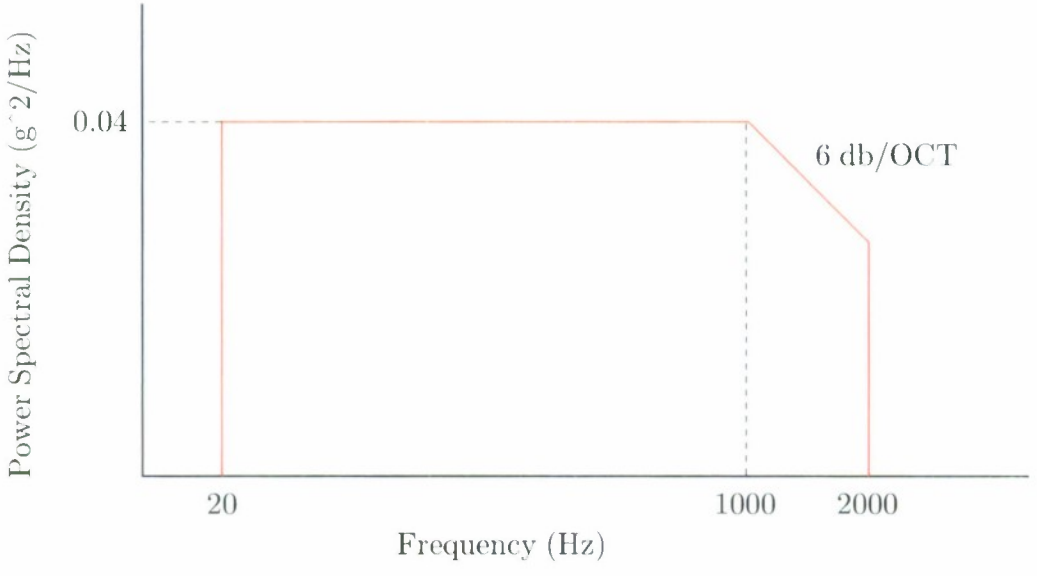


Figure 2.28: Input for the MIL-STD-810E minimum integrity test.

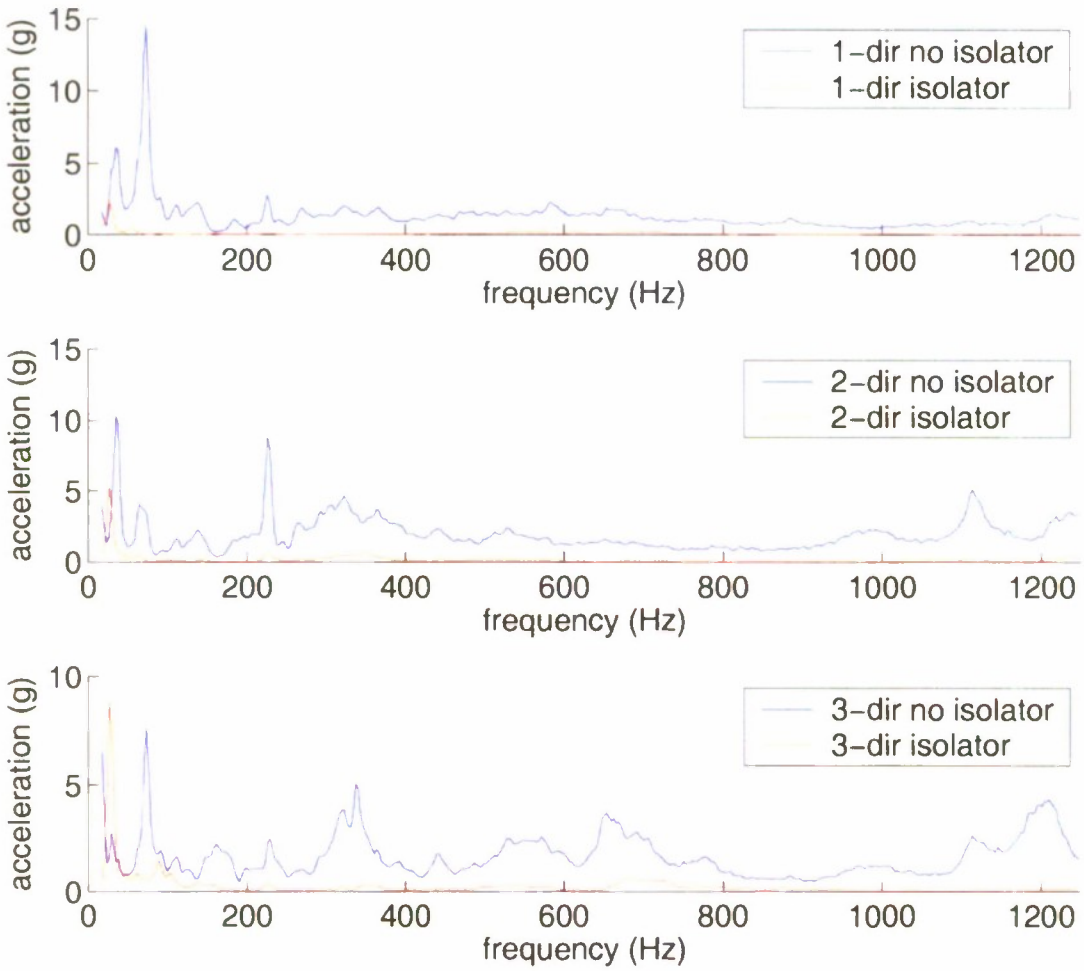


Figure 2.29: Acceleration results for the MIL-STD-810E minimum integrity test.

2.6 Experimental Response of Isolators

In addition to the full system, experimental tests were performed on the isolators individually. A suite of uniaxial tension/compression experimental tests for the isolator components were performed: see Figures 2.30-2.32 for the experimental setup.

Figure 2.30: Overview of a specimen in the Instron uniaxial testing machine.

The uniaxial tests were performed at three different loading rates to assess rate dependency: see Table 2.6 for the various tests performed and the corresponding results found in Figures 2.33-2.35.

Table 2.3: Summary of experimental uniaxial tests performed for a given strain rate.

Specimen	$\dot{\varepsilon} = 0.0072$	$\dot{\varepsilon} = 0.072$	$\dot{\varepsilon} = 0.72$
1	# 1a	# 1b	# 1c
2	# 2a	# 2b	# 2c
3	# 3a	# 3b	# 3c
4	# 4a	# 4b	# 4c

Figure 2.31: A specimen under uniaxial tension.

Figure 2.32: A specimen under uniaxial compression.

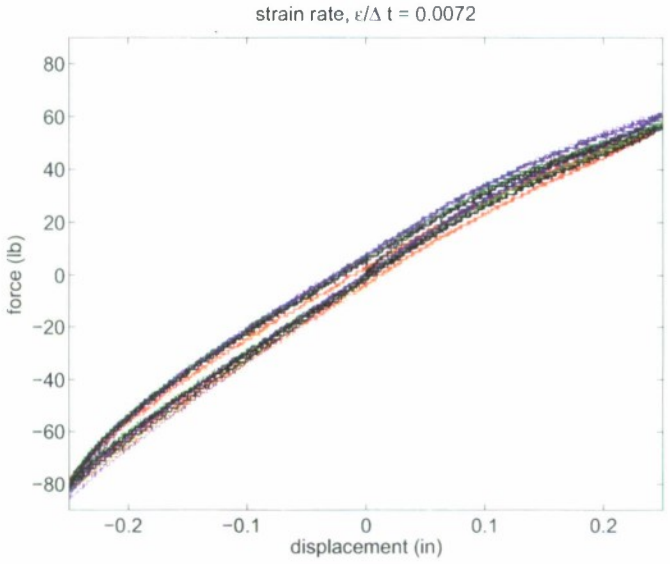


Figure 2.33: Load deflection curve for a strain rate of $\dot{\varepsilon} = 0.0072$.

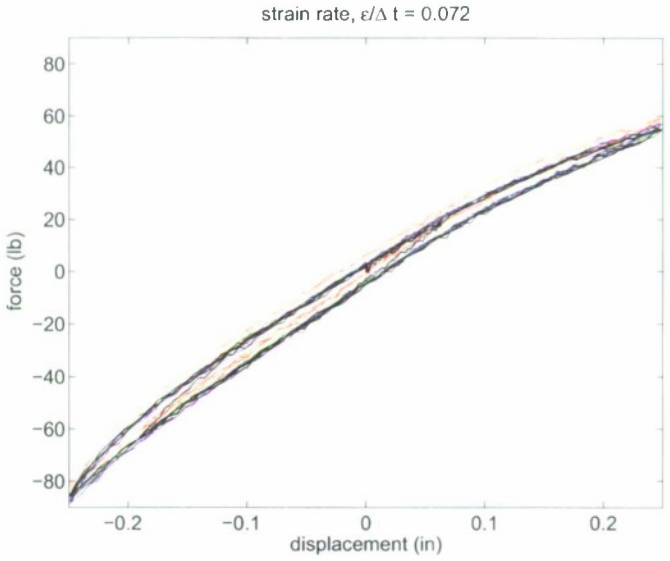


Figure 2.34: Load deflection curve for a strain rate of $\dot{\varepsilon} = 0.072$.

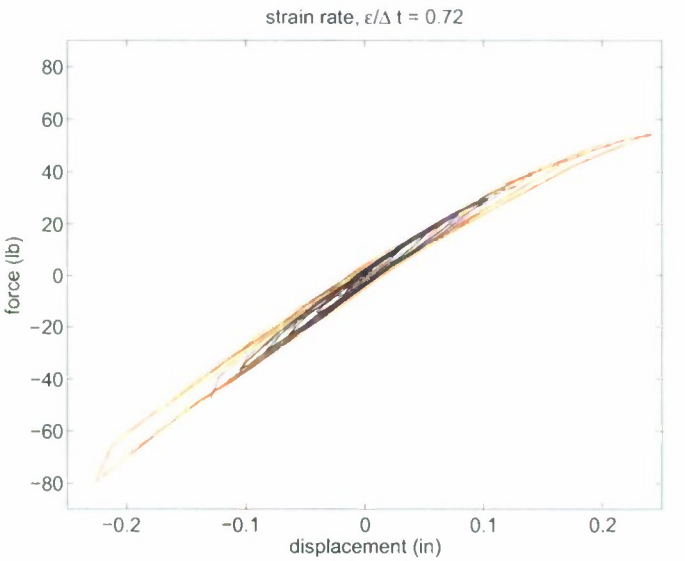


Figure 2.35: Load deflection curve for a strain rate of $\dot{\varepsilon} = 0.72$.

Section 3

Numerical Phase

3.1 Introduction

In this section the numerical model is presented. The model is comprised of the subassembly from the TRAP system, and the isolators which attach to the radar box. The steel components are modeled as presented in the previous discussion on the preliminary frequency analysis of the system. The primary revisions are the inclusion of a realistic model for the isolator, and the inclusion of damping as needed for realistic dynamic time history response simulations. Prior to presenting the improved model, the requisite mathematical basis of the model is summarized.

3.2 Mathematical Basis

3.2.1 Strong Form

Consider a three dimensional body, subjected to arbitrary loading, boundary, and initial conditions. The statement of the strong form of the initial

boundary value problem is written as:

$$(S)\begin{cases} \text{Find the function } u_i(x_i,t) \text{ such that} \\ \sigma_{ij,j} + b_i = \rho u_{i,tt} \quad \text{for all } x_i \in \Omega \times t \in \mathcal{I} \\ u_i = \overline{u}_i \quad \text{on } \Gamma_{ui} \\ \sigma_{ij} n_j = \bar{t}_i \quad \text{on } \Gamma_{ti} \end{cases} \tag{3.1}$$

where u_i is the displacement, σ_{ij} is the stress, b_i is the body force, ρ is the spatial density, $\overline{u}_i$ and $\bar{t}_i$ are the prescribed displacements and tractions on Γ_u and Γ_t, respectively, Ω is the spatial domain of the body, and $\mathcal{I}$ is the time domain.

3.2.2 Weak Form

The corresponding weak form consistent with the strong form is expressed as

$$(W)\begin{cases} \text{Find the function } u_i \in \mathcal{S} \text{ such that for all } v_i \in \mathcal{V} \\ \displaystyle\int_\Omega v_i \rho u_{i,tt}\, dV + \int_\Omega v_{i,j}\sigma_{ij}\, dV = \int_\Omega v_i b_i\, dV + \int_{\Gamma_t} v_i \bar{t}_i\, dS \end{cases} \tag{3.2}$$

where $\mathcal{S}$ and $\mathcal{V}$ are appropriate function spaces. To arrive at a form better suited for interpolation, the above equation is rewritten in the standard way, see [1], [2], and [3].

3.2.3 Finite Element Arrays

Substitution of an appropriate set of interpolation functions, $\mathbf{N}$, $\mathbf{B}$, based on the element class chosen, into the weak form yields the finite element arrays

$$\underbrace{\int_\Omega \mathbf{N}^T \rho \mathbf{N}\, dV}_{\mathbf{M}}\ \ddot{\mathbf{u}} + \underbrace{\int_\Omega \mathbf{N}^T \mathbf{C} \mathbf{N}\, dV}_{\mathbf{C}}\ \dot{\mathbf{u}} + \underbrace{\int_\Omega \mathbf{B}^T \mathbf{D} \mathbf{B}\, dV}_{\mathbf{K}}\ \mathbf{u} = \underbrace{\int_\Omega \mathbf{N}^T \mathbf{b}\, dV + \int_{\Gamma_t} \mathbf{N}^T \bar{\mathbf{t}}\, dS}_{\mathbf{F}}\ . \tag{3.3}$$

Note that the independent variables are typically interpolated in the same manner as the dependent variables (using $\mathbf{N}$), whereas the derivatives of the

dependent variable interpolation are denoted by **B**. Also the inclusion of viscous damping may also be introduced via Raleigh (classical) damping as

$$\mathbf{C} = \alpha\mathbf{M} + \beta\mathbf{K} \ . \tag{3.4}$$

Based on the topology of the system under consideration the above formulation may be reduced from a three dimensional continuum body to a three dimensional shell body in which the thickness direction is numerically integrated to reduce the dimensionality of the formulation. Shell theory is applicable for the majority of the geometry of the radar mount since individual parts comprising the body have one predominately small coordinate dimension relative to the other two coordinate dimensions, *i.e.* moderate thickness plates. See [4] and [5] for further elaboration on the limits of validity of this approach. With this assumption in place appropriate interpolations are implicitly chosen for the discretization through the element choice.

3.2.4 Finite Element Error Estimates

Since the interpolation functions used in the finite element arrays are commonly chosen from a family of polynomials, when implemented correctly, a linear response is reproducible for any patch of elements. Also since any smooth and continuous displacement field may be approximated by a collection of linear polynomials the solution can ideally be achieved to within a reasonable accuracy that will be consistent with the floating point system employed. The rate at which one approaches the solution under ideal conditions is, for a displacement based element, expressed as

$$||\mathbf{u} - \mathbf{u}^h|| \leq Ch^q \tag{3.5}$$

where q is dependent upon the order of the governing equation, h is the characteristic dimension of an element within the finite element discretization, and C is a problem dependent constant. Note the above estimate precludes the presence of any singularities.

To assess the convergence properties of the mesh discretization four consecutively finer meshes densities were considered. The elements chosen for the mesh discretizations were 4-noded quadrilateral elements for the shell surfaces and connector elements for the isolation mounts, see [6]. Although

not shown, a convergence study with respect to mesh density was performed to verify the meshed used in subsequent analyses.

3.3 Modal Dynamic Study

In the analysis of a system under external loads which vary over time, inertial and viscous forces can have an impact on the response depending upon the loading rate and frequency content. As such they must be accounted for. Consider the equations of motion for an finite degree-of-freedom system

$$\mathbf{M}\ddot{\mathbf{u}} + \mathbf{C}\dot{\mathbf{u}} + \mathbf{K}\mathbf{u} = \mathbf{p}(t) \tag{3.6}$$

where $\mathbf{M}$ is the mass matrix, $\mathbf{C}$ is the damping matrix, $\mathbf{K}$ is the stiffness matrix, $\ddot{\mathbf{u}}$ is the acceleration vector, $\dot{\mathbf{u}}$ is the velocity vector, $\mathbf{u}$ is the displacement vector, $\mathbf{p}$ is the forcing function vector, and the parameter n denotes the number of active degrees-of-freedom in the system. The natural frequencies of equation (3.6) are determined from the generalized eigenvalue problem

$$(\mathbf{K} - \lambda\mathbf{M})\,\boldsymbol{\phi} = \mathbf{0} \tag{3.7}$$

The natural circular frequencies of the system are simply $\omega_i = \sqrt{\lambda_i}$. If the frequency content and rate of the load is near or at one of the natural frequencies of the system then inertial and viscous forces contribute substantially to the response, *i.e.* dynamic response. Given the dynamic nature of the response the objective is to determine the motion of the n-coupled equations for a discrete system resulting from the finite element arrays (3.6), based upon an external forcing function and initial conditions. To determine the motion of (3.6) there are essentially two common techniques, namely a modal analysis method and a direction integration method.

In the development of a solution to the system of equations, (3.6), the equations of motion can be transformed into another set of coordinates (commonly called modal coordinates), which lead to a system of $n-$uncoupled equations. The $n-$uncoupled transformed equations are scalar single degree-of-freedom equations, which may be readily solved in closed-form for simple loading functions. Given the response for each individual degree-of-freedom the total response of the system is determined by summing up each individual response contribution. Since the solution is a summation of each

modal contribution the procedure is computationally efficient provided the number of modes chosen is a small subset of the total number of active degrees-of-freedom in the system. Although there are limitations to the modal procedure, the major ones being the inclusion of nonorthogonal or non-classical damping and nonlinearities (material or geometric). If the equations of motion cannot be efficiently uncoupled then an alternative to the modal analysis method is to numerically integrate the coupled equations of motion, which can be computationally expensive. There are a variety of procedures developed to numerically integrate systems of equations. The main one is the finite difference method, which approximates the derivatives through a difference approximation or stencil. Given the appropriate difference approximation a recursion relation is developed and a solution formed. It should be noted that the introduction of the difference equations into the system of equations introduces a finite discretization in time, in addition to the existing spatial discretization.

3.3.1 Modal Analysis

As alluded to in the previous section the response of the system may be expressed in terms of modal contributions, specifically the mode shapes multiplied by the modal amplitudes. The mode shapes are chosen such that one can sufficiently represent the response with a minimum number of terms.[1] The displacement response of a system can be expressed as

$$\mathbf{u}(\mathbf{x},t) = \mathbf{\Phi}(\mathbf{x})\mathbf{z}(t) = \sum_{i=1}^{m} \mathbf{u}_i = \sum_{i=1}^{m} \boldsymbol{\phi}_i z_i \tag{3.8}$$

where $\boldsymbol{\phi}_i$ is the i^{th} mode shape, z_i is called the i^{th} generalized or modal coordinate corresponding to the i^{th} mode, and m is the total number of modes to include in the response. Note the matrix of mode shapes, $\mathbf{\Phi}$, serves to transform the generalized or modal coordinate vector, $\mathbf{z}$, to the displacement response vector, $\mathbf{u}$. To enable the uncoupling of the equations of motion for an n-degree-of-freedom system the concept of modal decomposition or mode superposition is utilized.

The key concept behind modal superposition is to replace the displacement response $\mathbf{u}$ in the physical coordinate system by the generalized coordinate

[1] The mode shapes are analogous to trigonometric functions in a Fourier series.

vector and associated modal shape matrix, see equation (3.8). The orthogonality conditions of the modes shapes are then utilized to decouple the resulting equations of motion, now in terms of the generalized coordinate, z_i. A solution is then developed for z_i based on procedures developed for a single degree-of-freedom system. The final result is then transformed by $\boldsymbol{\phi}_i$ back to the physical space to yield $\mathbf{u}_i$. Summing the individual components $\mathbf{u}_i$ over the total number of modes in the system results in the total response for the system in physical coordinates. Typically only a relatively small subset (m) of the total modes n $(m < n)$ are needed to accurately represent the total response.

An overview of the procedure used to determine the response of the system, $\mathbf{u}$, is as follows:

- Solve the eigenvalue problem for the desired number of eigenvalues (square of the frequency), λ_i and eigenvectors (mode shapes) $\boldsymbol{\phi}_i$

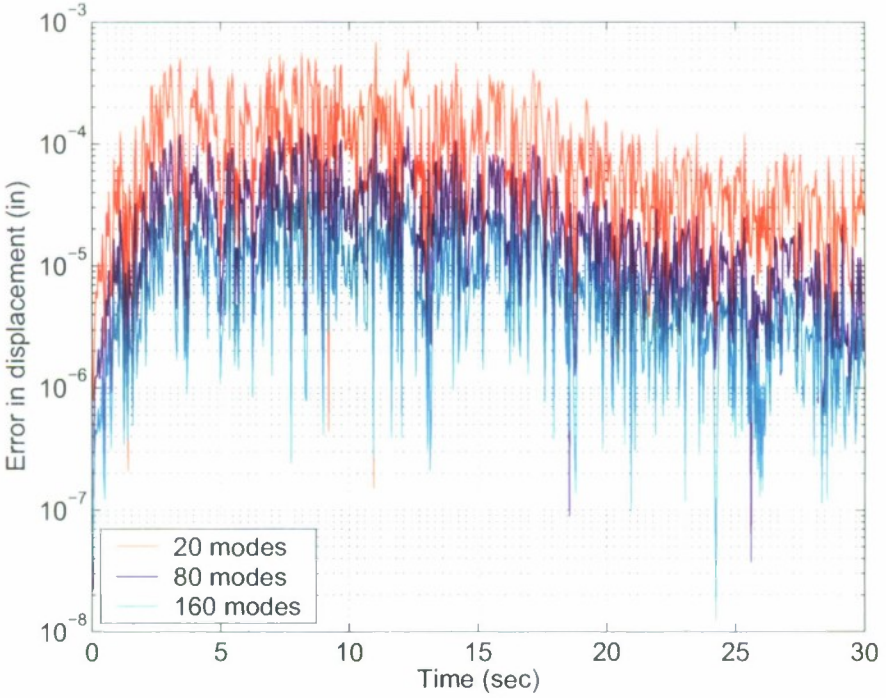
$$(\mathbf{K} - \lambda_i \mathbf{M})\, \boldsymbol{\phi}_i = \mathbf{0}$$

- Utilizing the results from the eigenvalue analysis determine the modal coordinate, z_i and perform a convergence study based on the number of modes, m, included in the summation

$$\mathbf{u}(\mathbf{x}, t) = \sum_{i=1}^{m} \mathbf{u}_i = \sum_{i=1}^{m} \boldsymbol{\phi}_i z_i$$

  Comparing the response for various number of modes will give a measure of the accuracy of the analysis with respect to modal contribution.

Figure 3.1 shows the error using 20, 80, and 160 modes with 320 modes as the baseline.



Error in time history response based on modal inclusion. Note 320 modes were used as a baseline to compute the error.

Figure 3.1: Convergence study on number of modes

3.4 Model for the Isolator

Given the experimental data determined from the experimental suite of uniaxial tension/compression tests a simple non-linear model (*i.e.* connector element) was developed which captures the macro behavior of the isolators. The finite element model used for comparison with the experimental data is shown in Figure 3.2, while the results of the finite element simulation in comparison to the experiments are shown in Figure 3.3.

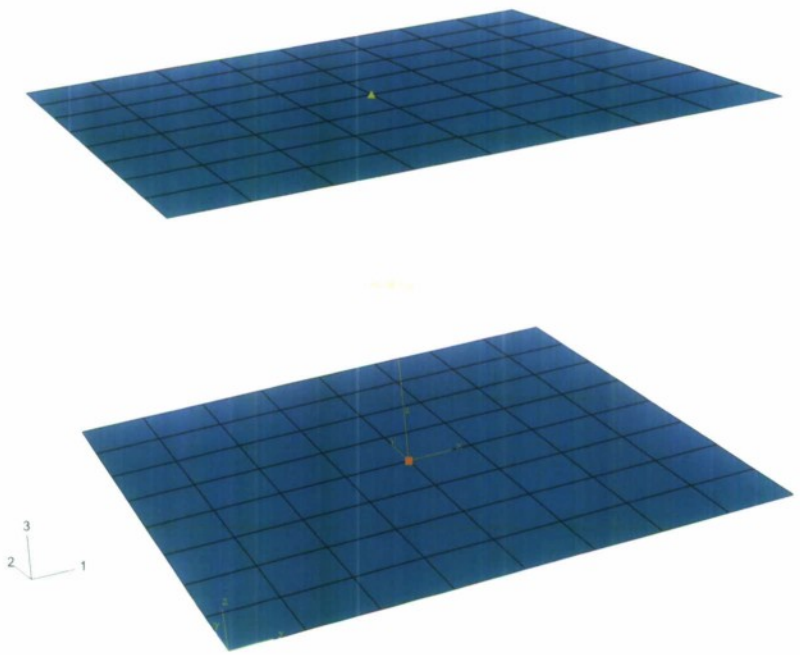


Figure 3.2: Finite element discretization of the connector model attached to two plates.

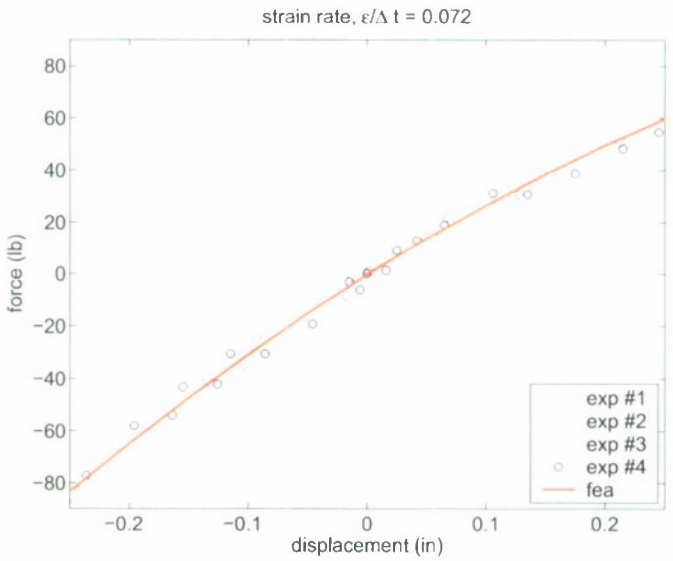


Figure 3.3: Load deflection curve for a strain rate of $\dot{\varepsilon} = 0.072$.

3.5 Eigenvalue Analysis

Utilizing the finite element model and the data from the experimental tests (*i.e.* damping) a series of numerical sine sweeps were performed to obtain the eigenvalues with and without the isolators. A summary of the numerical results for the eigenvalue analysis are found in the previous experimental section. A comparison of the numerical results with experimental data with and without isolators is shown in Tables 3.1-3.2.

Table 3.1: Comparison of eigenvalues for the TRAPS without isolators

mode	experimental (Hz)	numerical (Hz)	error
1	15.2	14.1	7.2%
2	28.6	25.5	10.8%
3	36.3	35.7	1.7%

Table 3.2: Comparison of eigenvalues for the TRAPS with isolators

mode	experimental (Hz)	numerical (Hz)	error
1	10.4	9.4	9.6%
2	16.5	14.2	13.9%
3	31.2	34.0	9.0%

Section 4

Conclusions and Recommendations

This section provides a brief overview of the findings of the study of the TRAPS radar mount.

4.1 Conclusions

The following items were concluded from the analysis as documented in the main body of the report.

- The isolation model was able to capture the response of the rubber bearing with reasonable accuracy for both the static response and dynamic response simulations.
- The system level model correctly captures the trends observed in the experimental data with regard to the natural frequencies of the system both with and without the proposed isolators.
- The sensitivity analysis lead to an understanding of:
  - the critical nature of connection detail modeling
  - the importance of boundary condition details to the problem

  - a rational approach to bounding behaviors in future design studies

- The work verified the correspondence of the numerical and experimental approaches thereby providing an outline for numerically testing alternative isolation mounts in the future without the need for expensive and time consuming experimental testing.

4.2 Closing Remarks

The conclusions listed above illustrate the fact that there is a great opportunity to improve the design-build-test cycle with regard to the mechanical performance of systems such as TRAPS. Through integration of numerical analysis, traditional design experience/tools, and the experimental program one can reasonably expect to vastly reduce the need for experimental testing during the product development stages.

Bibliography

[1] O.C. Zienkiewicz and R.L. Taylor. *The Finite Element Method: The Basis*, volume 1. Butterworth-Heinemann, Oxford, 5th edition, 2000.

[2] O.C. Zienkiewicz and R.L. Taylor. *The Finite Element Method: Solid Mechanics*, volume 2. Butterworth-Heinemann, Oxford, 5th edition, 2000.

[3] T.R.J. Hughes. *The Finite Element Method.* Prentice-Hall, New Jersey, 1987.

[4] S. Timoshenko and S. Woinowsky-Krieger. *Theory of Plates and Shells.* McGraw-Hill, New York, 2nd edition, 1987.

[5] A.C. Ugural. *Stresses in Plates and Shells.* McGraw-Hill, New York, 2nd edition, 1999.

[6] Hibbitt and Karlsson & Sorensen Inc. ABAQUS. http://www.hks.com.

Development of Reactive Polymer Coatings for Chemical Weapons Degradation

Project Investigator:

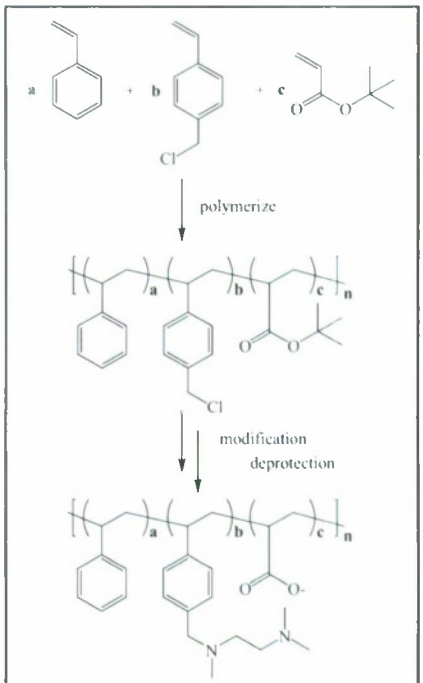
Chad Immoos
Chemistry and Biochemistry
California Polytechnic State University
San Luis Obispo, CA

Development of reactive polymer coatings for chemical weapons degradation

The purpose of this project is to develop new reactive polymer coatings capable of chemical warfare agent degradation. To achieve this goal we are synthesizing polymers and polyelectrolyte films containing reactive metal catalysts and/or metal nanoparticles and screening for reactivity against chemical warfare agent simulants. To facilitate our investigations we are developing analytical methods to monitor the degradation of CW agents in solution and in the surface immobilized polymers. Our first objective is directed towards the synthesis and characterization of reactive polymers containing transition metal or metal nanoparticle catalysts. The development of reactive polymer coatings for CW agent degradation would find utility on the battlefield as a passive method of equipment decontamination and as an additional layer of protection for the soldier. Our second objective is directed toward developing methods for evaluating the catalytic effectiveness of our catalyst-polymers. Initially, UV-visible spectroscopy will be used to screen the reactivity of our catalyst-polymers. The eventual goal is to monitor the adsorption and destruction of CW agents by these coatings using FTIR spectroscopy.

Description of work performed:

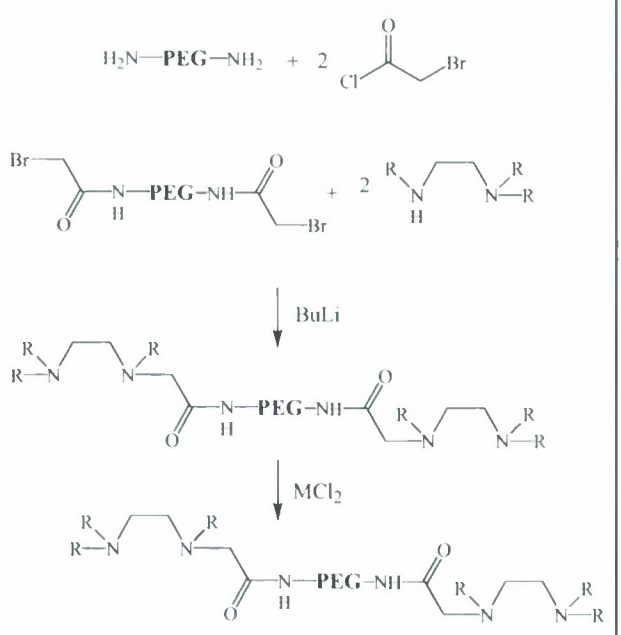
Synthesis of polymers containing transition metal catalysts. We are investigating the incorporation of transition metal catalysts into polymers using the method described by Menger and coworkers.[1] In particular, we have continued our studies of Cu catalyst-containing polystyrene. We have previously synthesized copolymers of polystyrene and chloromethylstyrene and modified them to contain chelating ligands capable of binding copper (II) ions. The resulting metal-containing polymer was demonstrated to effectively degrade Paraoxon, a chemical warfare agent stimulant. We have extended our investigations of catalyst-containing polymers by incorporating metal chelating ligands into traditional latex



Scheme 1. Synthesis of catalyst containing latexes.

polymers, according to Scheme 1. Synthetic efforts have created three different polymers with different ratios of chloromethylstyrene and *tert*-butylacyrlate. Attempts to modify the chloromethylstyrene to contain a metal chelate were unsuccessful. The resulting product appeared to be a highly cross-linked polymer gel. The polymer was insoluble in all solvents.

We are also investigating the incorporation of metal catalysts as small polymer additives. Initial work has focused on the synthesis of the metal chelate-containing polymer additive shown in Scheme 2. Our initial synthesis has been successful in creating a chelate-contained polyethylene glycol (PEG) compound. Incorporation of copper (or other metals) into the resulting modified PEG has been hampered by difficulties purifying the desired products. When these difficulties are overcome, the metal-containing PEG will be dissolved in organic solvents and added to other polymers, such as polystyrene, polyethylene, or polyacrylics. The polymer blends will be spread onto glass slides to analyze the catalytic activity of the resultant thin films. The transition metal loadings of the polymer blends can be controlled by varying the concentrations of metal-containing PEG.



Scheme 2: Synthesis of transition metal containing polyethylene glycol.

Immobilization of metal nanopartieles. Our previous studies have determined that MgO nanoparticles are effective at degrading chemical warfare agents. We have further investigated methods for the incorporation/immobilization of metal oxide nanoparticles on surfaces in a manner that retains this reactivity. We employed a method based on work reported by Rubner and coworkers for the deposition of metal oxide (SiO_2)

nanoparticles onto the surface of porous polyelectrolyte multilayer films.[2] We have generated porous multilayer films of polyacrylic acid (PAA) and poly(allylamine hydrochloride) (PAH) and incorporated MgO nanoparticles onto the surface. The nanoparticle-multilayer films have been tested for reactivity against Paraoxon and the surfaces characterized by atomic force microscopy (AFM). The nanoparticle-multilayer films show only modest reactivity toward Paraoxon; likely due to the method of deposition of the nanoparticles onto the multilayer films. We are currently changing our nanoparticle deposition conditions to remedy this problem. AFM images of the surfaces show that they are highly porous, Figure 1. This porosity increases the available surface area for immobilization of reactive nanoparticles.

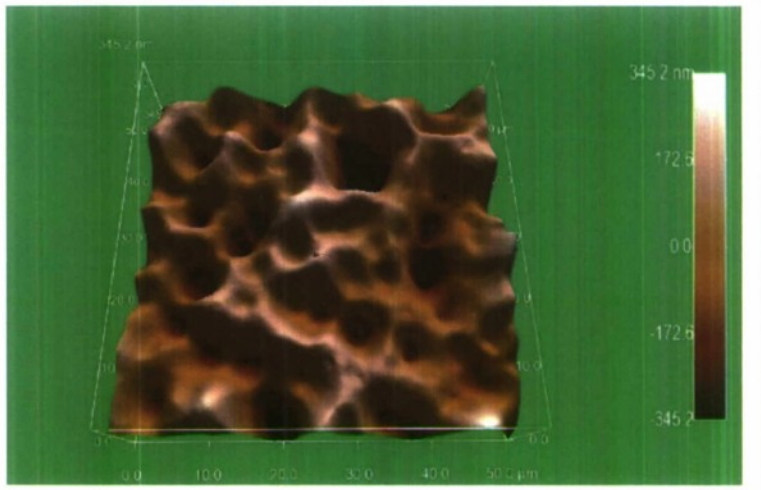


Figure 1: AFM demonstrating the pore size of the 100 bilayer MgO PAH/PAA porous slide. Pore size is estimated to be on average 5-10 µm wide. The scale is 50x50 µm.

Analysis of reactivity using UV-visible and infrared spectroscopy. Our initial studies have monitored the degradation of Paraoxon as a chemical warfare agent simulant. Paraoxon is an ideal CWA surrogate for development of reactive coatings and validation of our analytical techniques because it can be monitored by UV spectroscopy due to the strong aromatic absorption at 270 nm.[3] Loss of this absorption band and the appearance of an absorption band at 340 nm for the released nitrophenol can be used to monitor the degradation of Paraoxon in solution.

We have determined the reactivity of nanoparticle covered polyelectrolyte multilayer films using paraoxon. The multilayer film-coated slides were immersed in a cuvette containing paraoxon and a buffer. The reaction is most efficient at high pH and the buffer is included in the reaction solution to help control the pH of the solution. Addition of a MgO-coated PAA/PAH multilayer film slide containing 2 bilayers of MgO gives the following change in the UV-visible spectrum (Figure 2). The modest change in absorption is not surprising given the relatively small amount of MgO nanoparticles on the surface of the films.

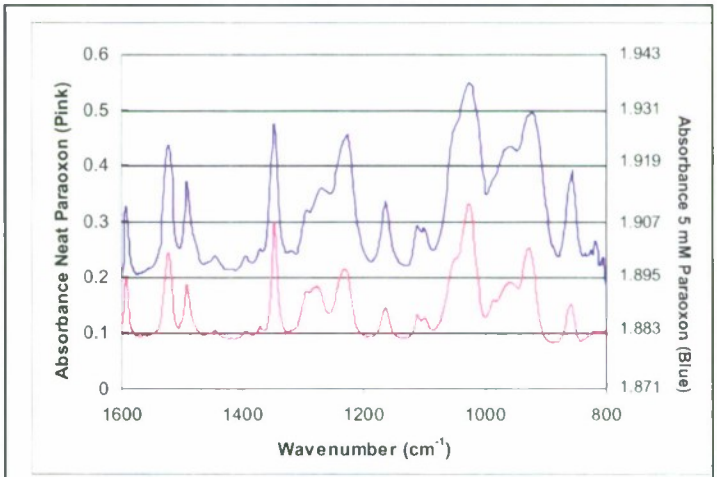


Figure 3: Comparison of the IR's for "neat" Paraoxon (pink) and 5 mM aqueous Paraoxon (blue).

In addition to UV-visible spectroscopy, we are currently investigating polymer films using infrared spectroscopy. These methods are still in development, but initial experiments have shown that the large water background present in the infrared region when working with dilute aqueous solutions of Paraoxon can be subtracted successfully, as shown in Figure 3. A "neat" sample of Paraoxon scanned with an air background on a fourier transform infrared spectrometer is compared to a 5 mM aqueous sample of Paraoxon (the concentration typically used for reactivity studies) scanned on a ZnSe crystal on an attenuated total reflectance (ATR) attachment with a water background. As one would expect the peak intensities are significantly different, but the number of peaks and wavelengths are nearly identical for the two samples. Currently, efforts are underway to determine if it is possible to track the hydrolysis of Paraoxon by infrared spectroscopy. The challenge with using an ATR-FTIR spectrometer is to not damage the ZnSe crystal, which is sensitive to basic solutions. Theoretically, there should be a shift in the peaks as more P-OH and aromatic C-OH groups are made. Depending on the sensitivity of this method, IR spectroscopy may replace UV-Vis spectroscopy as a method to monitor the hydrolysis of organophosphates. Because this method does not

rely on a colored product, use of the IR absorbances would be a tremendous step towards monitoring the degradation of other organophosphates that do not create colored products.

References

(1) Menger, F. M.; Tsuno, T. *J. Am. Chem. Soc.* **1989**, *111*, 4903-4907.
(2) Zhai, L.; Cebeci, F. C.; Cohen, R. E.; Rubner, M. F. "Stable superhydrophobic coatings from polyelectrolyte multilayers" *Nano Letters* **2004**, *4*, 1349-1353.
(3) Rajagopalan, S.; Koper, O.; Decker, S.; Klabunde, K. J. *Chem. Eur. J.* **2002**, *8*.

Toughening Mechanisms of Hybrid Fiber Reinforced Concrete Composites

Project Investigator:

Daniel C. Jansen
Civil and Environmental Engineering
California Polytechnic State University
San Luis Obispo, CA

INTRODUCTION

Conerete and other eement-based materials are weak in tension and brittle. For this reason engineers plaee steel reinforcing bars in struetural eonerete. More recent advanees with high-performanee eoneretes and eement-based eomposites have ineluded relatively short, diseontinuous fibers whieh increases the toughness, or ability to absorb energy, of the materials. The most eommon applieation of fibers in eonerete is the use of steel or polypropylene fibers in conerete slabs to help eontrol craeking. In eement-based eomposites, glass fibers are used in eement siding boards and roofing tiles; these eement based eomposites are tough enough that driving a nail through the produet will not ereate splitting or eraeking.

Fiber reinforeed eonerete and fiber reinforeed eement-based eomposites are becoming more and more eommon in construetion as the teehnology gains aeeeptanee, at the same time significant advanees are still being made. They are beginning to find their way into major struetural eomponents and into joints enhaneing the performanee of the strueture subjected to signifieant seismie or blast events. Furthermore, continuing advanees are being made with the introduetion of new fiber types and geometries and with the use of eombining fiber types, or hybrid fibers. With proper seleetion of hybrid fibers synergy ean be obtained; however, at this time only a few basie guidelines have been established regarding whieh fibers will be synergistie eombinations.

The outeomes of this proposed work has direet relevance to both an area of foeus of the ONR/DOD and to the teehnology park initiative. While the direet purpose of this researeh is to gain a better understanding of the initiation and growth of eraeks in hybrid eement-based eomposites and the synergy of different fiber types in their ability to prevent craeks from initiating and then growing into size eraeks. The outeomes of this work will aid in the better understanding of the behavior of fiber reinforeed eement-based eomposites and will aid in the seleetion of fibers whieh interaet synergistieally.

OBJECTIVES

The objeetives of this work inelude:

1. development of testing protoeols with existing equipment at Cal Poly to perform toughness testing of fiber reinforeed eonerete in eompression and flexure.
2. development of software for interpreting eomputer vision data for the visualization and quantification of cracks in concrete.
3. shed some light onto the initiation and development of eraeks in eement-based eomposites eontaining hybrid fibers.
4. eompare results of the standard flexural toughness test for fiber reinforeed eonerete (ASTM C1609) with quantified eraek propagation from direet tensile tests.
5. determine whether hybrid fiber eombinations ean synergistieally improve the toughness of fiber reinforeed eoncrete.

Ultimately, this work eould allow us to better model the behavior of fiber reinforeed eement based eomposites for the purpose of designing better performing struetures or blast mitigating elements and help us understand the interaetions of individual fiber types in hybrid composites so we ean better prediets whieh fiber eombinations will work synergistieally.

BACKGROUND

There are two general eategories of fiber sizes: miero and macro. Maero fibers are the most common type used in eonerete; they are typically about 2" (50 mm) long and 0.04" (1 mm) in diameter. Maero-fibers are good for keeping large eraeks from developing and for absorbing energy (improving toughness); however, they do little in preventing miero-eracks (10 to 25 μm wide) from developing. Conversely, miero-fibers are usually used in eement-based eomposites sueh as eement siding boards or roofing tiles. These fibers are exeellent at preventing eraeks from initiating and ean inerease the tensile strength (first eraeking strength) of eement based materials. Miero-fibers, however, break onee eraeks reaeh approximately 250 μm (0.0001") in width, subsequently they do little to absorb energy or improve post-craeking energy. So for optimal response of a fiber reinforeed eonerete where one has both high tensile strength and high toughness, one should be eombing fiber types (hybrid fibers).

Despite the number of papers published and the amount of studies performed on hybrid fibers, there seems to be no eonsensus on what exaetly produees optimal eombinations of fibers capable of produeing synergy in performanee. A number of the papers have demonstrated a distinet synergy between fibers, more eombinations of fibers fail to be synergistie. While the work being proposed will not establish "rules" for determining optimal fiber eombinations, it will provide some insight to the eombined effeets of different fibers and how eraeks initiate and grow.

A seeond item whieh eomplieates evaluation of results is the standard test method. In the United States, ASTM C1018 (1997) was the original standard test method for determining the toughness of fiber reinforeed eonerete; however, with this method, determination of the load at craeking was highly operator dependent, so results from this test method varied greatly. ASTM C1399 (2004) was subsequently added in 2000 to determine the residual strength of fiber reinforeed eonerete. While ASTM C1399 provided reproduеible results, the method only gave how mueh stress the eonerete eould withstand just after eraeking, and had no bearing on the true toughness of the eonerete. In 2005, a new ASTM standard, ASTM C1609 (2005), was added to replaee ASTM C1018. In Japan and in Europe, eompletely different standards are used, most of which have some flaw either in the sensitivity of the test, operator dependenee, or seatter in the results (Banthia and Soleimani, 2005). Trying to eompare results from these widely varying and un-reproduеible test methods has further eomplieated eomparisons of results.

Another issue with the standard toughness testing with third-point loaded beams is the test does not properly eharaeterize the type of eraeks whieh form. Multiple eracking with small spaeing is favorable to one wide eraek. These two seenarios eould result in the same load versus defleetion eurve for beams as long as the beam with the single wide eraek has fibers whieh can bridge a wide eraek. Some researehers still prefer to perform uniaxial tension tests, although they are more diffieult to perform, on fiber reinforeed eonerete sinee fiber types whieh result in multiple eraeks have duetilities whieh are measurably greater than materials where the damage loealizes into one eraek.

EXPERIMENTAL PROGRAM

Materials

Fibers

This report outlines the performance of two different batches, both with multiple fibers. The first batch contains both steel and carbon fibers, and the second batch contains both steel and alkali-resistant glass fibers. Table 1 below provides a list of the fiber properties present in these two batches.

Table 1: Fiber Properties

Manufacturer	Product Name	Material	Length (mm)	Diameter (microns)	Aspect Ratio	Density (lb/ft^3)	Specific Gravity	Modulus of Elasticity (ksi)	Tensile Strength (ksi)	Description
Bekaert Corp.	Dramix 65/35	Steel	35	550	64	490	7.85	29,000	160	Deformed Filament
Saint-Gobain	Anti-Crak HP	AR Glass	19	14	--	167.2	2.68	10,000	250	Monofilament Fiber
Hexcel	Magnamite	Carbon	6.4	--	--	111.7	1.79	33,500	655	Chopped Fiber

Mortar Matrix

The fiber-reinforced mortar used in this study was made from a consistent mix design comprised of sand (passing a No. 4 sieve), cement, fly ash, and water. The mixture proportions are provided below in Table 2 in both per batch and per cubic yard basis. Fibers of various types were then placed in the wet mix and allowed time to disperse to ensure a random distribution.

Table 2: Mortar Mix Design

Material	Specific Gravity	Amount per Batch	Amount per yd^3	
		As Made Wt.(lbs) Agg. OD state	As Made Wt. (lbs) Agg. SSD state	As Made V. (yd^3) Agg. SSD state
Sand (passing #4)	2.53	97	1,690	0.396
Cement	3.15	60	1,027	0.193
Fly Ash	1.94	20	342	0.104
Water	1.00	32	518	0.307
TOTAL		209	3,578	1.000

Experimental Protocols

Batching & Casting

The sand used in this study was in the oven dry state prior to mixing. The sand was placed in the mixer first, along with an adequate amount of water to ensure the sand was fully saturated. Portions of the cement were then added to the mixer, along with some of the remaining water. A similar process was followed for the addition of the fly ash. Once all of the cementitious materials and water were added to the mixer, the mortar was allowed to mix for a few minutes before any fibers were added. Once the fibers were placed in the mixer, they were allowed several minutes to disperse. Super-plasticizer was often added to the mixer to promote workability of the mix, since higher fiber contents tend to impede workability. The fiber and air contents of the two batches examined herein can be found in Table 3 below.

Table 3: Fiber Type and Content and Air Content Present in Batches

Property	Batch 1		Batch 2	
	Fiber 1	Fiber 2	Fiber 1	Fiber 2
Fiber Type	Steel	AR Glass	Steel	Carbon
Fiber Manufacturer	Bekaert	Saint-Gobain	Bekaert	Hexcel
Fiber Name	Dramix 65/35	Anti-Crak HP	Dramix 65/35	Magnamite
Fiber Length (mm)	35	19	35	6
Fiber Density (pcf)	489	167	489	112
Volume %	1.38%	0.31%	1.65%	0.36%
Air Content %	11.46%		14.02%	

Three different types of specimens utilized in this study included tension (dog-bone shape), compression cylinder, and flexural beam. Casting of the specimens was completed by scooping the finished mix into one of the forms, then placing the form on a vibration table until the amount of air bubbles emerging from the concrete had significantly decreased. Figure 1 shows the three types of specimens, in their respective forms, after casting.

Figure 1: Specimens in Forms after Casting

Curing

Once removed from the forms, the specimens were placed in a curing environment which consisted of either a submerged lime bath, or a humid curing room. The specimens were allowed to cure until a few days prior to testing.

The tension specimens had to be adequately dried in order to be painted with the speckle pattern which allows the computer vision system to track specimen displacements. The ends of the cylinders also had to be dry in order to place the sulfur capping compound on either end. The flexural beams did not need to be dry for any additional preparations, but were removed to ensure that all specimen types were kept in similar environments.

Compression Tests

The compression test conducted in this study were performed on 4" diameter by 8" tall cylindrical specimens. The specimens were cast in plastic cylindrical molds and capped with a sulfur compound a day prior to testing.

The average diameter was recorded so that the cross-sectional area, and subsequently the stress of the specimen could be determined. The capped specimens were placed in a compressometer with an LVDT to measure the strain as shown in Figure 2.

Figure 2: Capped Compression Specimen with Compressometer Fixture.

Flexural Toughness Tests

The flexural test utilized in this study is that specified by ASTM standard test C 1609-05, with 4" x 4" x 16" beam specimens. Flexural beam specimens were cast in a vertical mold to provide four smooth, cast faces. The specimens were placed horizontally on two supports, spaced 12 inches apart, and the load was applied at a spacing of 4 inches, centered between the supports as shown in Figure 3.

A pair of frames were constructed which attach to the specimen directly above either support at mid-height of the specimen. A cross-member runs between the frames, from which the midpoint deflection of the specimen is recorded by a pair of linear variable differential transformers (LVDTs). As the ends of the specimen displace, the frame moves with it, thus eliminating the need to subtract out the deflection occurring at the ends. A comparison of the measured mid-span deflection versus the machine reported deflection is shown in Figure 4. After the test is complete, the specimen dimensions at the location of failure are recorded in order to compute the stresses and modulus of elasticity.

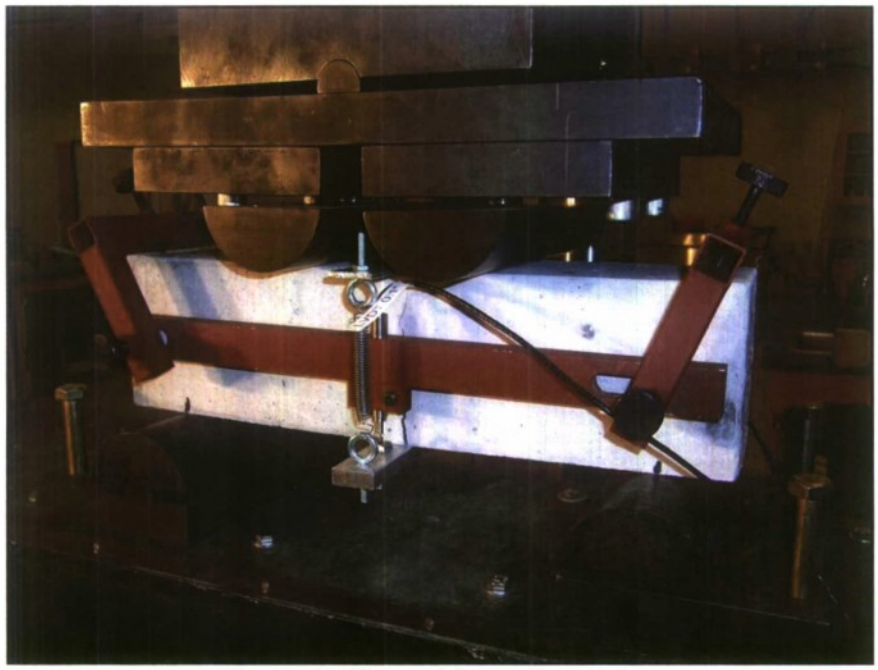


Figure 3: Flexural Beam Test Setup

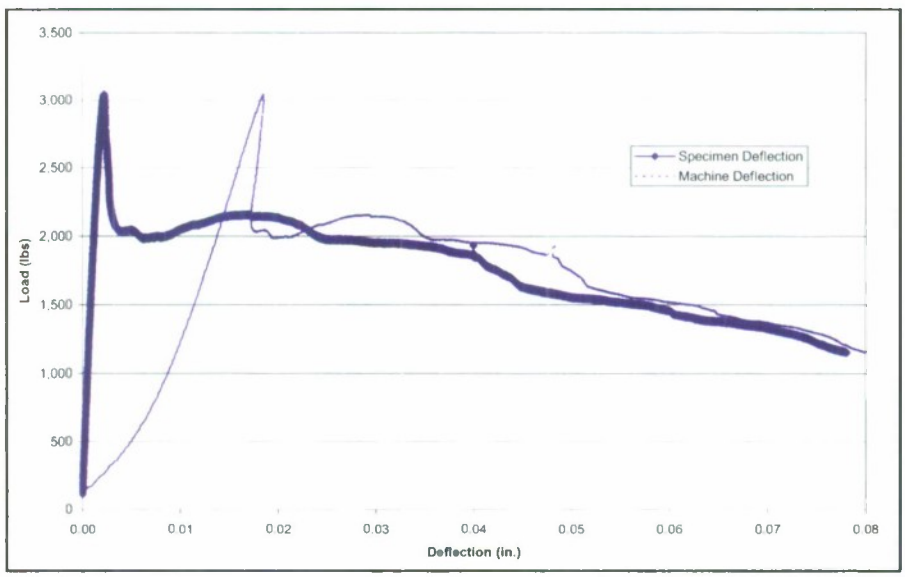


Figure 4: Comparison of Specimen and Machine Deflection for Flexural Beam Test

Tensile Tests

The tension samples were held in the test machine by a pair of steel plates at both the top and bottom of the specimen. A thin piece of neoprene rubber was inserted between the specimen and the plate to improve grip and help protect against any stress concentrations that might result from an irregular or bumpy specimen surface. The steel plates were bolted tightly together around the

specimen to provide adequate friction to develop the loads necessary to cause specimen failure. Specimen dimensions are provided in the drawing below in Figure 5 (a). A photograph is provided in Figure 5 (b) demonstrating the tension test setup.

Control of the test was provided by the average strain as measured by two extensometers mounted to either side of the specimen. The average strain was set to a constant rate and the computer controlling the testing machine determined the corresponding force needed to impart the stated strain. Specimen dimensions at the failure location are recorded in order to compute the stress and modulus of elasticity of the specimen.

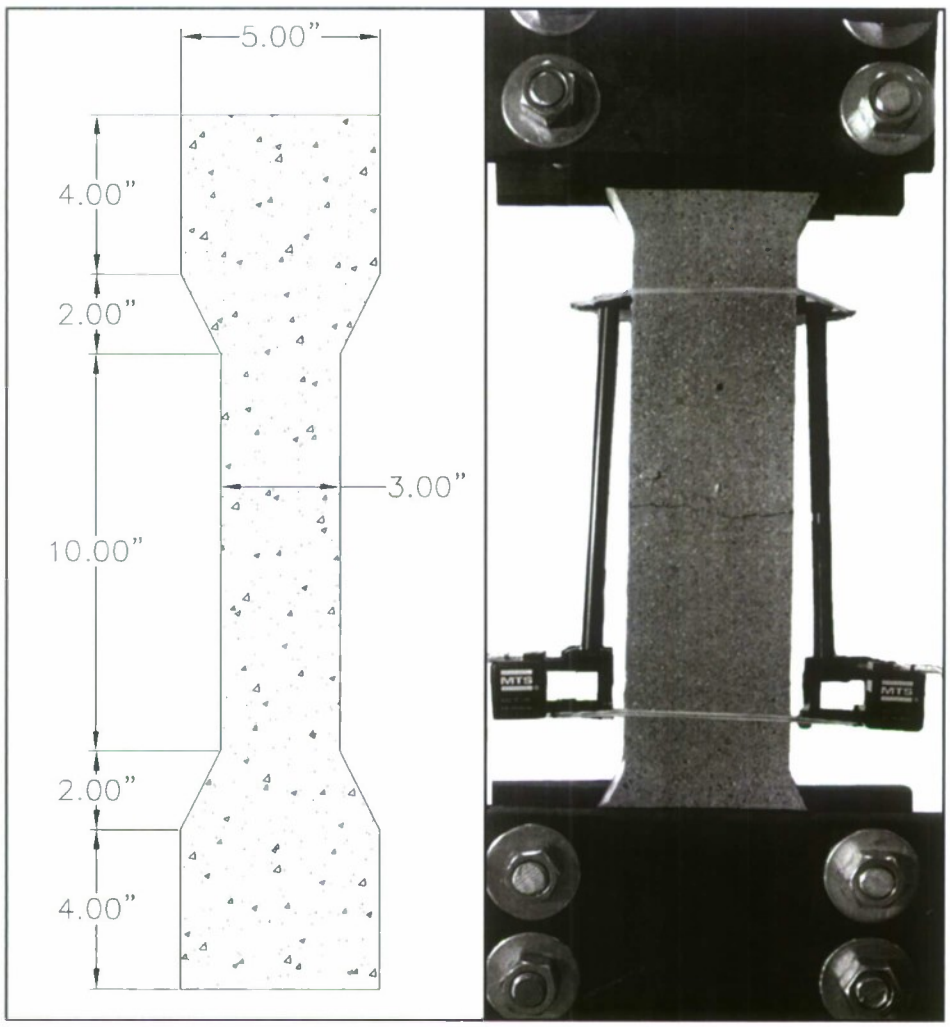


Figure 5: **(a)** Tension Specimen Dimensions, and **(b)** Specimen in Machine Fixture.

Computer Vision

One major facet of the research project was the application of digital image analysis to measure the formation and growth of cracks in the specimens during testing. Digital image analysis, or computer vision as it is often called, is a method of measuring relative displacements across the surface of a specimen.

Test Preparation and Data Collection

Preparation of the fiber-reinforced concrete samples involved the application of a uniform speckle pattern to each side of the sample with various shades of colored spray paint. The size of the pattern is controlled to ensure that the data analysis software will produce reliable results; a pattern that is either too fine or too coarse may result in erroneous data.

Image data is collected during tensile tests using two high-resolution digital cameras (8 and 10 megapixel) via remote shooting. The cameras are positioned on either side of the test specimen and are configured to capture images simultaneously every 10-20 seconds. Before testing began reference and calibration images are taken for later analyses. During each test up to 400 images were captured (up to 200 pictures per camera).

Image Correlation

The principle upon which computer vision is based is the matching of image subsets between two different images. The first image, or the reference image, is of the un-deformed specimen (i.e. prior to the application of any load). The second image, or the target image, is of the deformed specimen at some point during loading.

The computer vision system determines the change in location of an image subset, or sub-image, between the reference image and the target image. This change in location is given in terms of a relative horizontal and vertical displacements. To determine the new location of the sub-image, a comparison is made between a sub-image on the reference image and an arbitrary sub-image on the target image, until the best match on the target image is found. Once the new location of the sub-image is found, the vector displacement is determined as shown in Figure 6. Matching of image subsets is then performed at a regularly spaced grid of nodes (each node being at the center of a sub-image) across the entire image. Matches between sub-images on the reference image and the target image by calculating the normalized image cross-correlation coefficient given in Equation 1:

$$r(m,n) = \frac{N\sum_i\sum_j TR - \sum_i\sum_j T\sum_i\sum_j R}{\left[N\sum_i\sum_j [T]^2 - \left[\sum_i\sum_j T\right]^2\right]^{1/2}\left[N\sum_i\sum_j [R]^2 - \left[\sum_i\sum_j R\right]^2\right]^{1/2}} \quad (1)$$

In this equation, $r(m,n)$ is the normalized image cross correlation coefficient and R and T are the intensity, or brightness, values of pixels within the sub-images in the reference and target image, respectively. This calculation gives a measure of how well a sub-image from the reference image and a sub-image from the target image match up. The algorithm in the image analysis program attempts to minimize the normalized cross correlation coefficient for each two sub-images being compared. When the best match has been found, a vector (m,n) is given. This is the vector displacement of the sub-image from the reference to the target image. Sub-pixel resolution is obtained using interpolation between pixels.

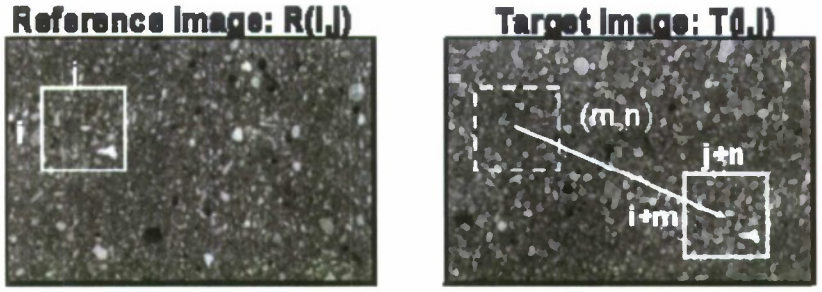


Figure 6: Sub Image Mapping

For the purpose of this project, commercially available software called VIC-2D was used to perform the image correlation; however, the results from VIC-2D are given as arrays of numbers with only the coordinates (x and y) and the displacements (u and v) given in units of pixels for each subimage analyzed. To eliminate poor data and to locate and quantify cracks in concrete, software was developed as part of this project.

Software Development for Crack Quantification

Additional analysis of the VIC-2D output was necessary to obtain useful characteristics of the tested fiber-reinforced concrete samples. A graphical user interface was written with Microsoft Excel in the Visual Basic (VBA) programming language. The primary function of the interface is to aid the user in easily determining when, where and to what extent cracks form in the concrete. Options for analysis include the following: rotation of the given reference frame to account for circumstances such as a tilted camera, user-defined origin location, subtraction of rigid body displacement, and a user-defined threshold for invalid data. Figure 7 shows the main screen of the user interface.

All data is compiled and manipulated into the state defined by the user through the interface. Data is reduced from the pre-existing conditions used by Vic-2D, such as coordinate system orientation and measurement units, into a more intuitive data structure. Using a calibration image, the coordinates and displacements are converted from pixels into mm or inches.

The conditions necessary to constitute a crack are determined based on the noise of the system. A crack that is smaller than the noise of the system cannot be conclusively detected. The current and typical testing conditions suggest that the noise of the system is such that a crack of 10 microns can be detected. Therefore this threshold is used in calculating the presence of a crack, although it may be redefined as testing conditions change. A change in displacement (delta u or v, depending on the orientation of the test) greater than the threshold suggests a crack has occurred. A linear regression of the displacement versus location plot is the final step in characterizing a crack. The regression is performed on the curve in the general location of the crack and results in the width and specific location of the crack. This process is repeated for each image collected during testing, giving overall cracking characteristics of the concrete to include when and where a crack started and how and where it propagated.

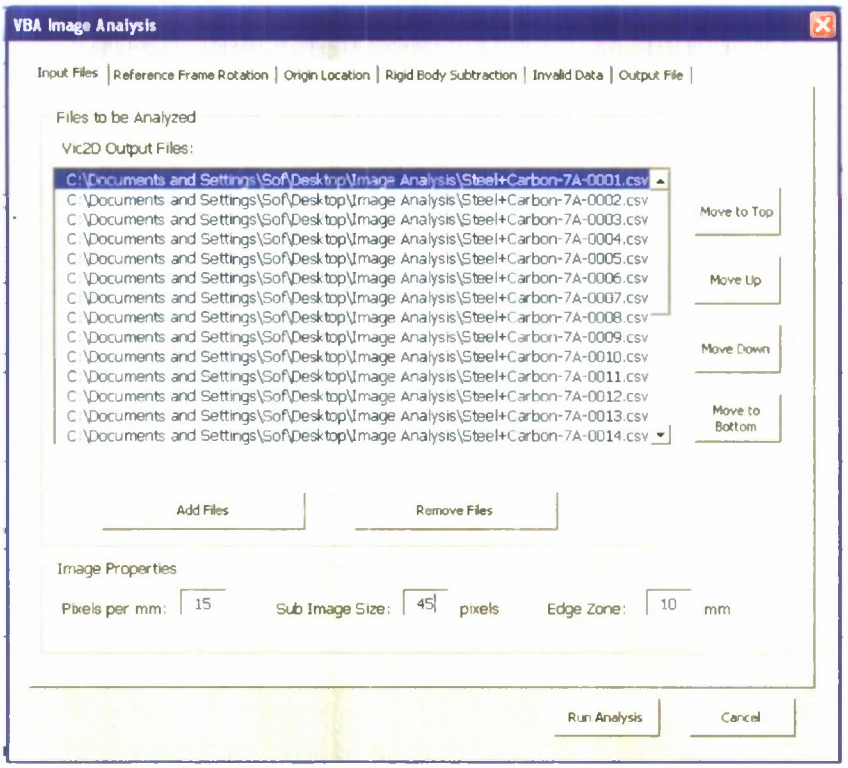


Figure 7: VBA Image Analysis User Interface

Results and Conclusions

Compression Test Results

The compression tests for the two batches show considerably different results from one another, but are fairly consistent within batch. Batch 2, in particular, demonstrated highly repeatable results in regards to strength with a total variance of roughly 9%. Table 4, shown below, presents the strength and modulus, and predicts a coefficient for the relationship between the modulus of elasticity and square root of the compressive strength (Equation 2).

Table 4: Comparison of Compressive Specimen Properties

	Specimen	Avg Dia. (in)	Area (in^2)	Load (lbs)	f'_c (psi)	Avg f'_c (psi)	E (psi) *	Avg E (psi)	SQRT (f'_c)	E / SQRT (f'_c)
Batch 1	Cylinder 1	4.019	12.686	53,409	4,210	4,579	1,843,498	1,857,183	64.88	28,412
	Cylinder 3	4.012	12.642	65,387	5,172		1,915,011		71.92	26,628
	Cylinder 4	4.016	12.667	55,166	4,355		1,813,039		65.99	27,473
									Avg =	27,504

	Specimen	Avg Dia. (in)	Area (in^2)	Load (lbs)	f'_c (psi)	Avg f'_c (psi)	E (psi) *	Avg E (psi)	SQRT (f'_c)	E / SQRT (f'_c)
Batch 2	Cylinder 1	4	12.566	30,003	2,388	2,470	1,095,062	1,095,674	48.86	22,411
	Cylinder 2	4.012	12.642	32,760	2,591		1,072,295		50.91	21,064
	Cylinder 3	4.016	12.667	31,994	2,526		1,068,291		50.26	21,257
	Cylinder 4	4.014	12.654	30,074	2,377		1,147,048		48.75	23,529
									Avg =	22,065

* NOTE: Modulus of Elasticity calculated from 5% of maximum stress to 50% of maximum stress.

$$E_c = X\sqrt{f_c'} \quad (2)$$

where: E_c = modulus of elasticity [psi]
f_c' = compressive strength [psi]
X = coefficient

Batch 2 shows an average compressive strength of 2,470 psi, or roughly 54% of that found for Batch 1. The modulus of elasticity displayed a similar trend between the two batches, with average values of 1,857,183 psi and 1,095,674 for Batch 1 and 2, respectively. For plain concrete specimens, the coefficient in the modulus of elasticity equation has been found to be approximately 57,000. The average coefficients found in this study are roughly half that, with values of 27,504 for Batch 1 and 22,065 for Batch 2. A graph is presented below, in Figure 8, which illustrates these trends. Aside from Specimen 3 in Batch 1, the other specimens tended towards the same residual load at larger strains. One explanation for the reduced strength using the carbon fibers in Batch 2 could be related to the increased air content from 11% to 14% between batches 1 and 2; although the 3% increase in air content typically would not decrease the strength as much. Subsequently, more work should be conducted to investigate to explain why the carbon fibers result in such poor strengths.

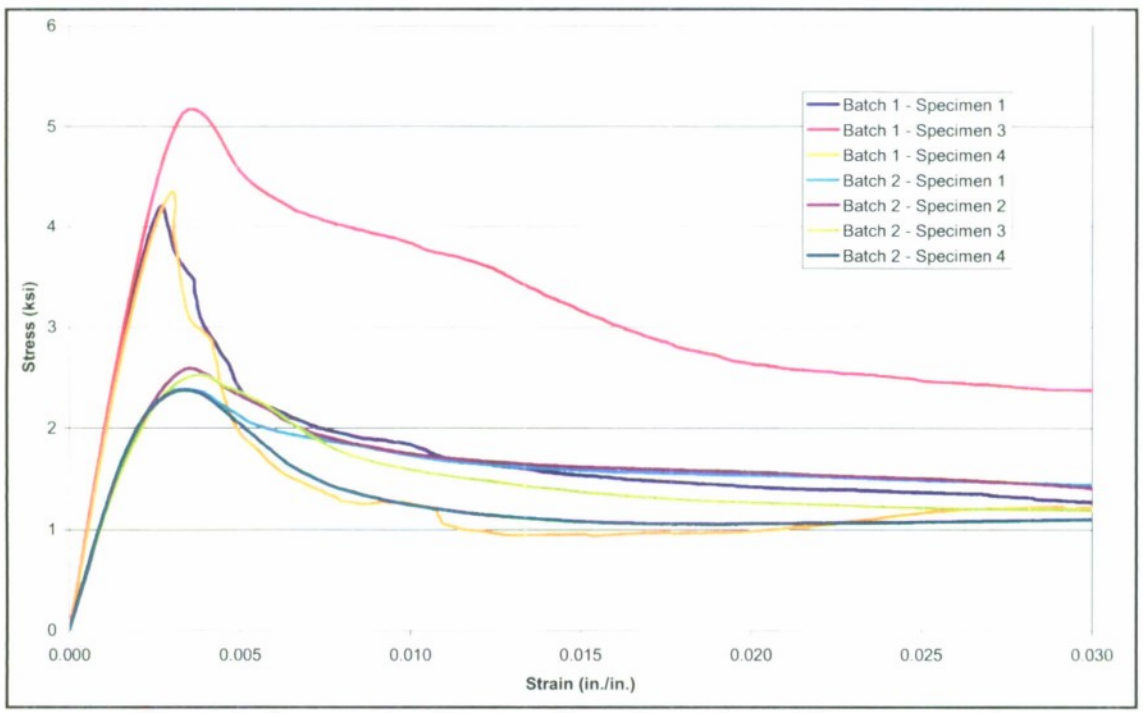


Figure 8: Compressive Stress vs. Strain Comparison of Batches 1 and 2

Flexural Toughness Test Results

The flexural beam test yielded results that were a bit closer together between Batch 1 and Batch 2, as shown in Table 5. The toughness was once again calculated by finding the area beneath the load versus deflection curve.

The bending tensile strength, also known as the modulus of rupture (f_r), as computed by Equation 3, varied by roughly 20% from Batch 1 to Batch 2, as opposed to nearly 100% for the tensile test. The modulus of elasticity was also much closer between batches, varying by roughly 30%. A comparison between the two test methods cannot be made for the toughness since the tensile toughness for Batch 1 could not accurately be determined. A diagram of the flexural stresses is included in Figure 9.

Table 5: Flexural Specimen Properties and Results

	Specimen	Width (in)	Depth (in)	Load (lbs)	f_r (psi)	Avg f_r (psi)	E (psi) *	Avg E (psi)	Toughness**	Avg Tough
Batch 1	Beam 1	3.917	4.010	3,047	581	559	2,320,631	2,190,565	136.1	145.9
	Beam 2	3.907	4.010	2,933	560		2,176,020		215.5	
	Beam 3	3.917	4.011	2,817	536		2,075,042		86.1	

	Sample	Width (in)	Depth (in)	Load (lbs)	f_r (psi)	Avg f_r (psi)	E (psi) *	Avg E (psi)	Toughness**	Avg Tough
Batch 2	Beam 2	3.907	4.014	2,584	493	457	1,580,813	1,587,128	81.6	88.5
	Beam 3	3.912	4.012	2,209	421		1,593,443		95.4	

* NOTE: Modulus of Elasticity calculated from 20% of maximum load to 60% of maximum load.

** NOTE: Toughness calculated at a deflection of 0.08 inches.

$$f_r = \frac{PL}{bd^2} \tag{3}$$

where: f_r = modulus of rupture [psi] P = peak load [lbs]
L = span length [in.] b = specimen width [in.]
d = specimen depth [in.]

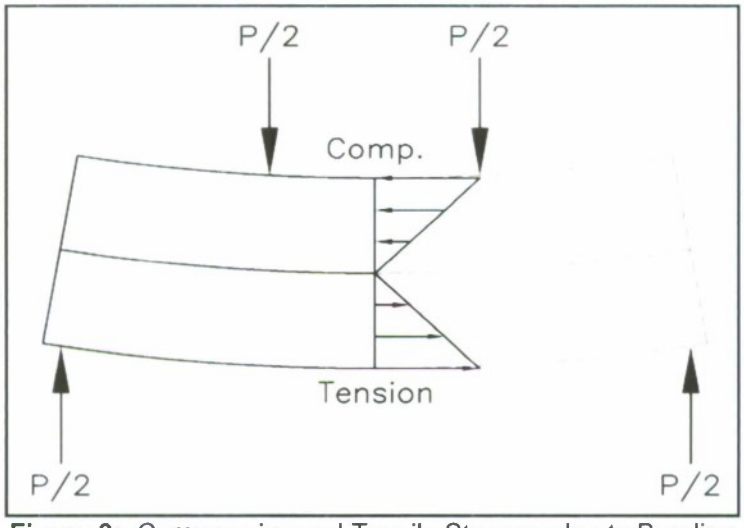


Figure 9: Compressive and Tensile Stresses due to Bending

The flexural beam test achieved fairly consistent peak loads and elastic moduli within a single batch, but residual loads and toughness are shown to be far less dependable. Figure 10 shows flexural tests for both batches.

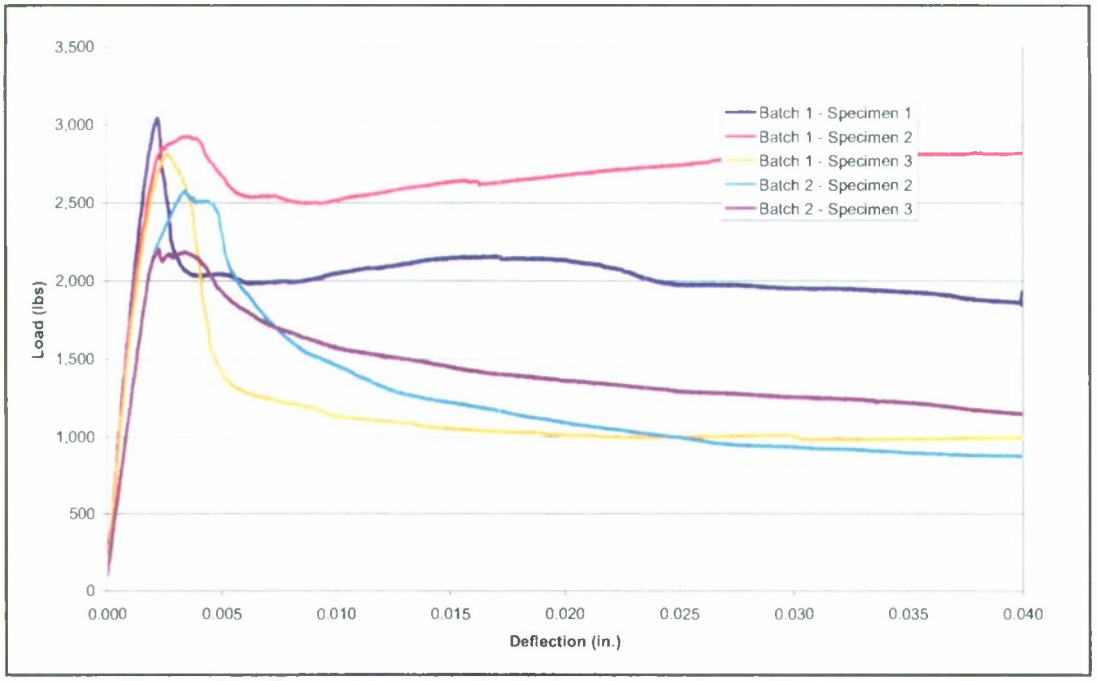


Figure 10: Flexural Load vs. Deflection Comparison of Batches 1 and 2

Tension Tests

The tension tests run on the two batches produced results that were quite different from one another. Table 6, provided below, presents ultimate tensile stresses, tensile moduli, and tensile toughness for the two batches. The toughness was calculated by finding the area beneath the load versus deflection curve and is a way of measuring the specimen's ability to absorb energy.

Batch 1 exhibited much larger load carrying ability as compared to Batch 2, evidenced by the average ultimate tensile stresses of 398 psi versus 202 psi, as noted below. The modulus of elasticity also varied considerably, with Batch 1 yielding an average modulus of 3,016,726 psi, and Batch 2 showing a modulus of 2,152,198 psi. The tensile toughness was unable to be accurately calculated for Batch 1, since the load dropped suddenly from tensile to compressive just after achieving peak tensile load. The cause for this is unknown, as this behavior has not been observed in any of the previous tension tests. The toughness is more likely to be influenced by flaws in the material, such as inadequate reinforcement or air voids. Toughness depends highly on the post-cracked behavior of the material, whereas the strength and modulus of elasticity are computed from the pre-cracked condition of the specimen. Thus the toughness can be highly variable from batch to batch, or even within a single batch, as indicated by the variability seen in Batch 2.

Figure 11 shows the stress versus strain test results of the tension specimens from Batches 1 and 2 illustrating the differences in both the ultimate and residual load carrying capacity of the specimens.

Table 6: Tensile Specimen Properties and Results

	Specimen	Width (in)	Thickness (in)	Load (lbs)	f_t (psi)	Avg f_t (psi)	E (psi) *	Avg E (psi)	Toughness**	Avg Tough.
Batch 1	Form 2	3.007	2.006	2,117	351	398	3,064,709	3,016,726	--	--
	Form 3	3.006	1.936	2,644	454		3,164,011		--	
	Form 4	3.006	1.963	2,295	389		2,821,459		--	

	Specimen	Width (in)	Thickness (in)	Load (lbs)	f_t (psi)	Avg f_t (psi)	E (psi) *	Avg E (psi)	Toughness**	Avg Tough.
Batch 2	Form 6	3.000	1.946	1,403	240	202	2,051,994	2,152,198	68.6	54.5
	Form 7	3.002	1.935	1,236	213		2,314,571		71.9	
	Form 9	3.008	1.899	872	153		2,090,028		23.2	

* NOTE: Modulus of Elasticity calculated from 20% of maximum stress to 60% of maximum stress.

** NOTE: Toughness calculated at a strain of 1.5%.

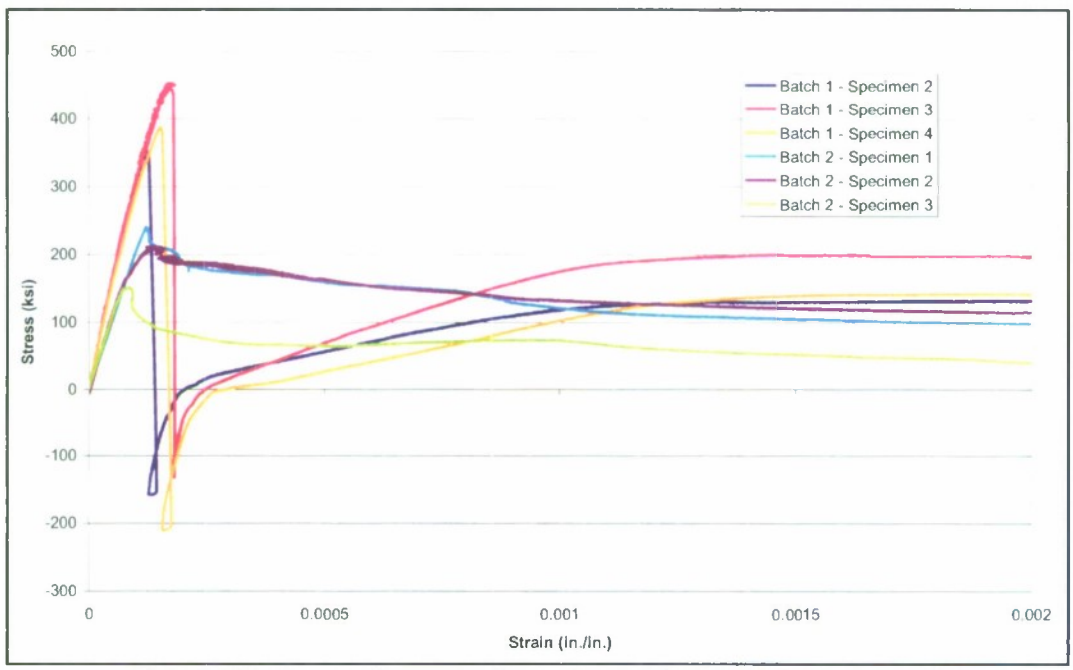


Figure 11: Tensile Stress vs. Strain Comparison of Batches 1 and 2

Computer Vision Analysis of Batch 2

The following are the results of the analysis of tensile testing on a fiber reinforced concrete of one sample from Bateh 2 (1.65% steel fiber and 0.36% carbon fiber). A rectangular AOI was selected in Vic-2D and a subset and step size of 45 and 10 pixels respectively were specified. Multiple cracks occurred until image 50, at which point existing cracks further developed, therefore full results are only shown for the first 50 images.

Two distinct cracks occurred during this test. Analysis of the sample indicated that the first eraek oeeurred at image 23, while the second crack formed at image 40. Figure 12 below shows the tensile stress vs. strain curve for the sample with the first occurrences of cracks indieated. The start time of the test was 0.00 seconds; therefore the appearanee of the first crack oeeurred 333.5 seeonds (5 minutes, 33.5 seeonds) after beginning the test. Likewise, the appearance of the second crack oeeurred at 558.5 seconds (9 minutes, 18.5 seconds).

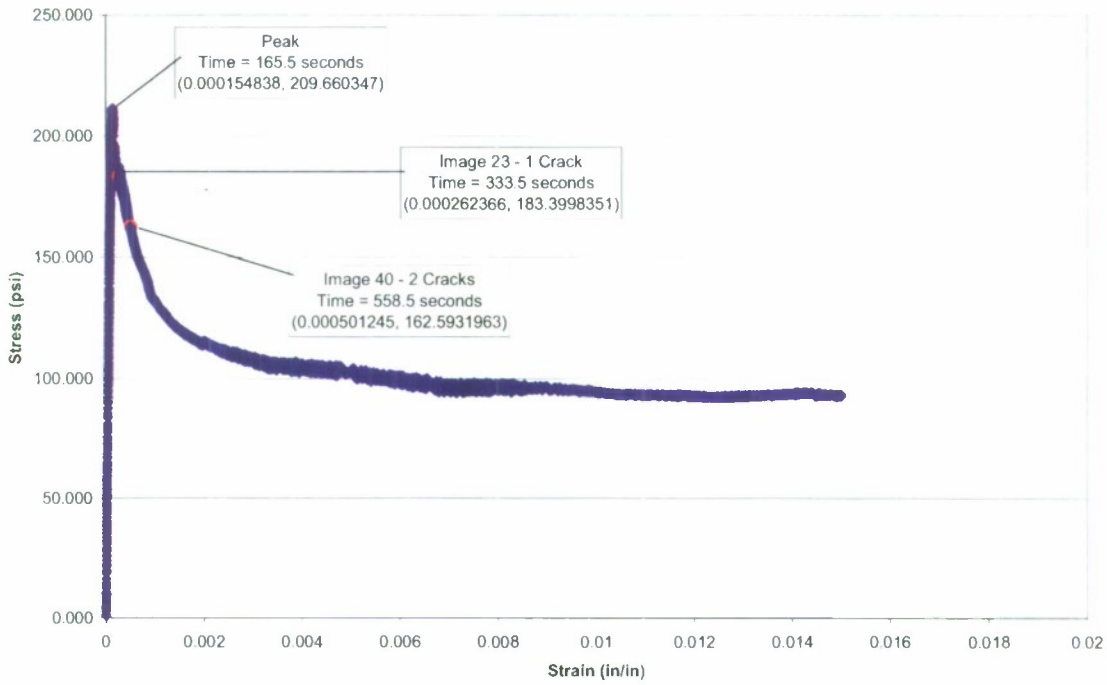


Figure 12: Batch 2 Sample 1 Stress-Strain Curve and Images

Image 23 and 40 are investigated more closely and Figures 13 and 14 show the displacement in the y-direction along the x-coordinate where cracks are observed. The discontinuities in the above figures represent cracks. It can be seen that the crack present in image 23 is clearly present in image 40, and another independent crack has also formed. The width of the cracks were determined using linear regression. Figure 14 shows the linear regression performed on the two cracks in image 40.

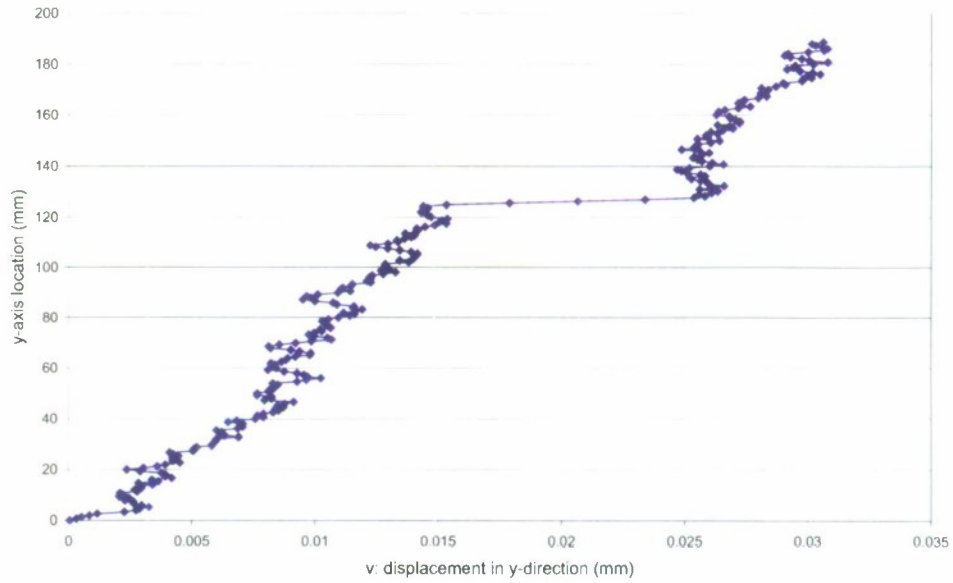


Figure 13: *y*-displacement (*v*) along *x* = 0.67 mm for Image 23

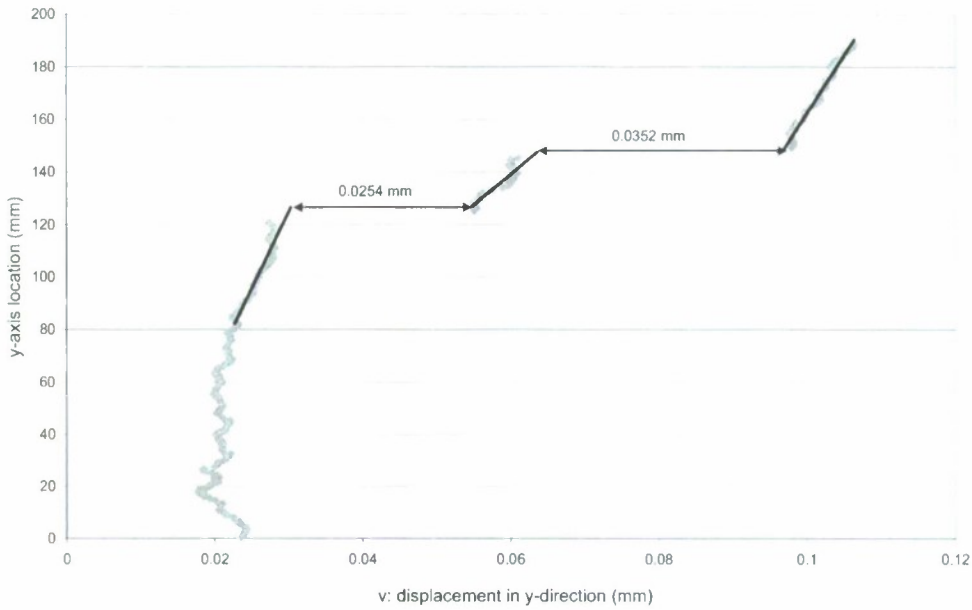


Figure 14: *y*-displacement (v) along x = 50.67 mm for Image 40

As indicated in the figure, there are two cracks in image 40, one of which is 0.0254 mm (25.4 μm) wide and the other is 0.0352 mm (35.2 μm) wide.

Contour plots for each of these images were created to emphasize the findings of the image analysis procedure. Figure 15 shows the beginning of the crack in image 23. The tight horizontal lines indicate a large relative displacement, which suggests a crack.

Similarly, Figure 16 shows the cracks present in image 40. The highlighted portion of the specimen shows the location of multiple cracking along one x-coordinate. The propagation of the cracks is also evident as the move from the outside of the specimen to the inside.

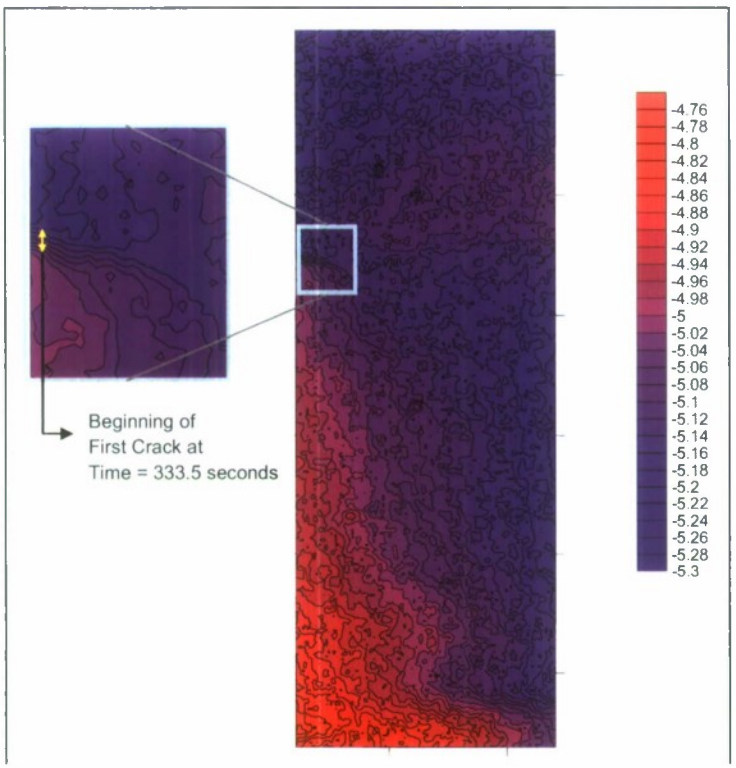


Figure 15: Image 23 Contour Plot

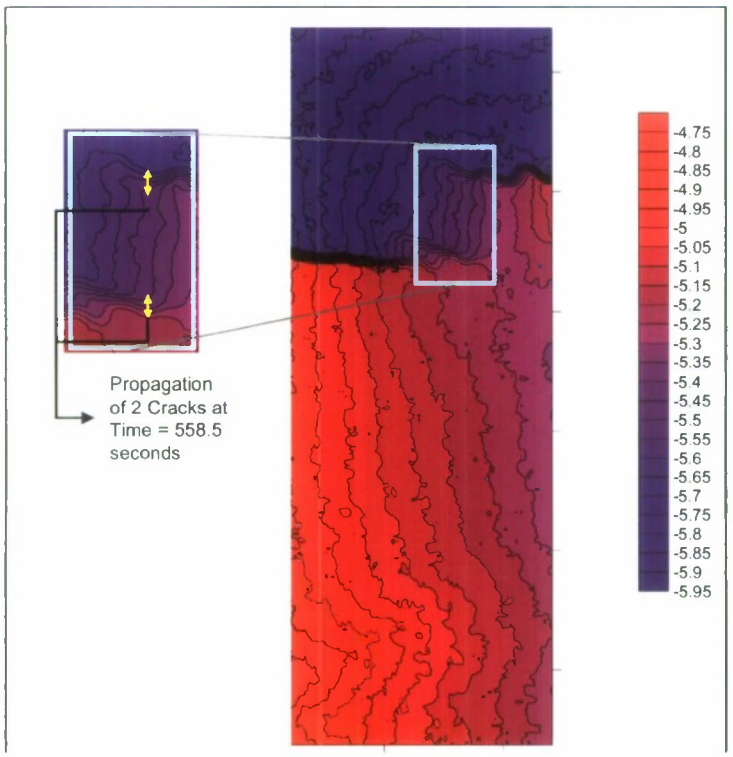


Figure 16: Image 40 Contour Plot

OUTCOMES

The work presented was for the development of test methods and investigation into hybrid fiber reinforced concrete composites. The work demonstrated the testing procedures and the developed software, although the data and conclusions were very limited. The primary contribution to the PI and his students has been the leap forward in the tests and analysis available for conducting future work. The PI has already used the results in support of the related proposal "High Performance Hybrid Fiber Reinforced Concrete for Bridge Decks" submitted in response to a Caltrans RFP for $297,643. A second proposal is intended for submission to NSF to continue the work with this project. The PI and additional students are still continuing the work to produce more sound results which should lead to results of publishable quality.

REFERENCES

ASTM C1018 (1997). "Standard Test Method for Flexural Toughness and First Crack Strength of Fiber Reinforced Concrete (Using Beam with Third-Point Loading)," ASTM International, West Conshohocken, PA, 7 pp.

ASTM C1399 (2004). "Test Method for Obtaining Average Residual-Strength of Fiber Reinforced Concrete, ASTM International, West Conshohocken, PA, 6 pp.

ASTM C1609 (2005). "Standard Test Method for Flexural Performance of Fiber Reinforced Concrete (Using Beam with Third-Point Loading)," ASTM International, West Conshohocken, PA, 8 pp.

Banthia, N., and Nandakumar, N. (2003). "Crack Growth Resistance of Hybrid Fiber Reinforced Cement Composites," *Cement and Concrete Composites*, Vol. 25, pp. 3-9.

Banthia, N., and Soleimani, S.M. (2005). "Flexural Response of Hybrid Fiber-Reinforced Cementitious Composites," *ACI Materials Journal*, Vol. 102, No. 6, pp. 382-389.

Cohen, I.E. (2003). "Tension Stiffening and Cracking in Fiber-Reinforced Concrete with FRP Reinforcement," Masters' Thesis, Tufts University, Medford, MA, 355 pp.

Cronquist, D.S. (2005). "Size Effect in Deep Beams Designed with the Strut and Tie Method," Masters' Thesis, California Polytechnic State University, San Luis Obispo, 204 pp.

Shiek, M.T. (2005). "The Influence of Specimen Geometry on Tensile Strength as Determined by the Brazilian Splitting-Tension Test – A Three Dimensional View," Masters' Thesis, California Polytechnic State University, San Luis Obispo, 146 pp.

Proteomic Approach to Evaluation of Dairy Components for Improvement of Human Performance

Project Investigators:

Rafael Jimenez-Flores and Michael Black
Dairy Science
California Polytechnic State University
San Luis Obispo, CA

Industry Collaborators

Dairy Management, Inc.

The original objectives of this project were:
Objectives:

- To characterize the biologically active proteins in MFGM using a proteomic approach.
- To assess the effect of processing on the presence and activity of MFGM using commercial conditions.
- To evaluate the biological activity of the MFGM components using microbial and immunological methods.

Achievements toward objectives:
To characterize the biologically active proteins in MFGM using a proteomic approach.

Specific tasks within objectives.

- *To characterize the biologically active proteins in MFGM using a proteomic approach.*

Task 1. Optimize separation of the minor proteome of the MFGM and standardize identification of the proteins by MALDI-TOF.

Results. This task proves to be more difficult that predicted and during the first six months we had unacceptable gels (see figures below). In response, we took a proteomics course at the UC Davis Proteomics facility. Towards the last quarter of this project we had a successful 2D Gel, and the comparison to published data is extremely good. We are currently proceeding with Objective 2 below. The figure of successful 2D gel with our preparation of MFGM proteins will allow us to maintain a program in this area of research.

In addition to this task, while we were optimizing the conditions for membrane proteins analysis in the 2D PAGE, we could analyze some of the proteins' activity by analyzing immunoglobulins by ELISA (see figure below).

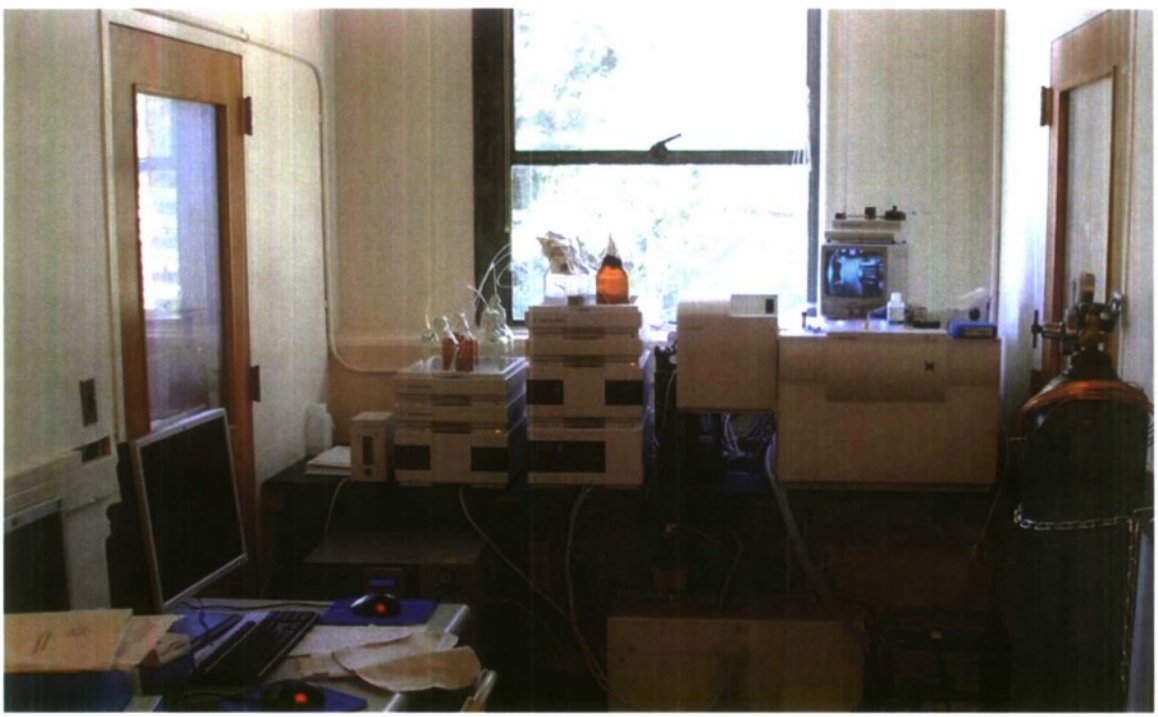
Task 2. Verify identity of the proteins by comparative studies using proteomics.

Although we did not get to complete Objective 1, we had some progress towards increasing our knowledge on the biologically active proteins of the MFGM. Using Western Blott techniques and Human TLR4 antibodies, we could identify the presence of this important factor for passive immunity in the milk membrane. We will use this valuable information to follow the presence and potential cross-linking of this protein with other soluble proteins in milk.

In previous experiments we could use the proteomic approach to find out with detail what protein and which segment of the proteins were interacting. With some effort it was also identified the precise amino acids involved in the reaction. That level of analysis needed the knowledge of the sequence of amino acids in each of the proteins. In this case, it may be necessary a more extensive search since it is not clear that TLR 4 is the precise reactive protein in the cow's milk. However, it has been reported that cross species

similarity exists, and it is possible that we can get much more information from our current experiments.

Seminal to this task was also the purchase and installation of our HPLC-MS system. Part of the research time was spent finding out the optimal system and its location. We had to submit for approval several proposals and gain authorization for remodeling a room to locate the instrument. This activity took place while performing the tasks proposed.

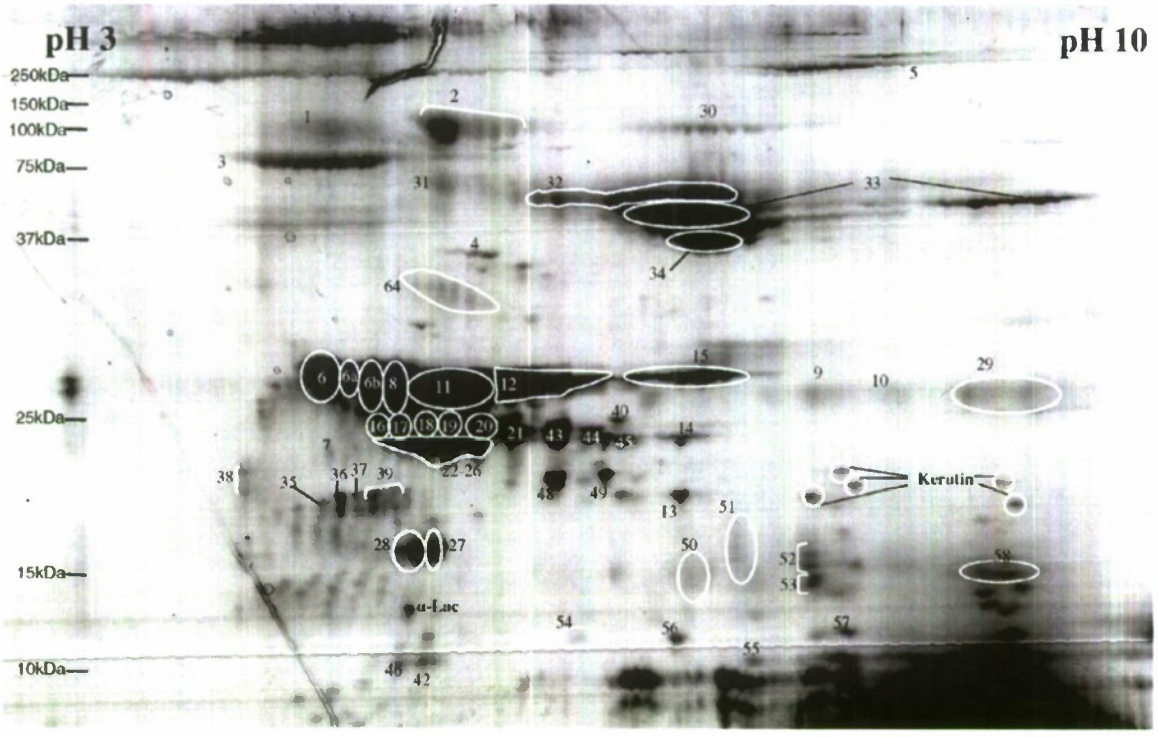


- *To evaluate the biological activity of the MFGM components using microbial and immunological methods.*

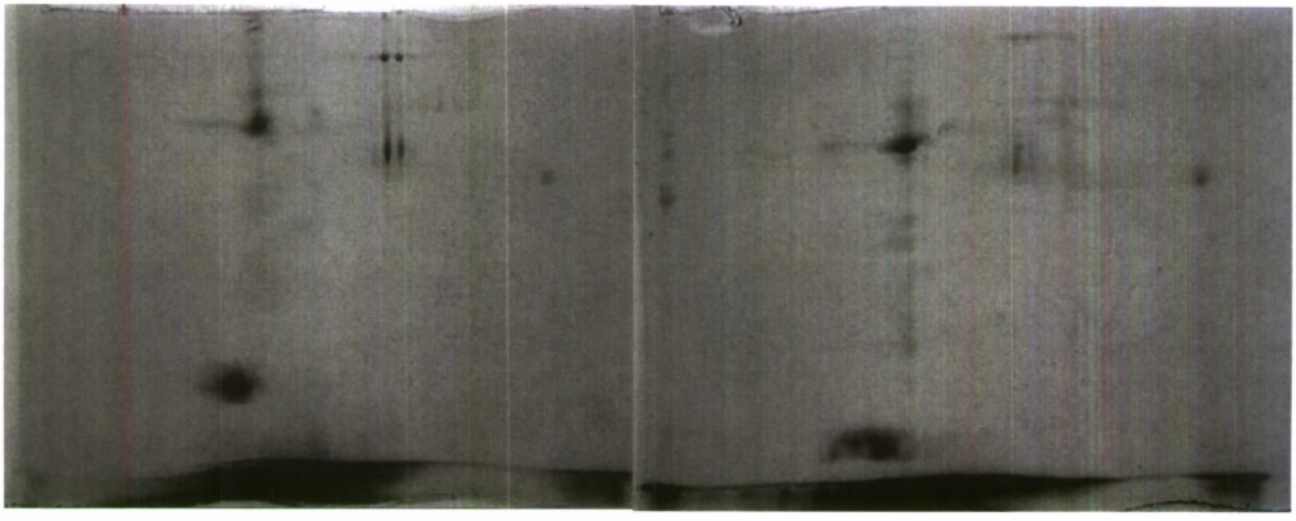
  Task 5. Measure the influence of MFGM fractions, before and after proeessing, on the proteome of probiotie baeteria

We accomplished some of the fundamental steps towards this task. We did not complete the proteomie evaluation of the components in milk. However, there was significant progress on the eontrol experiments that indieate the degree of denaturation of biologically active proteins in milk. The figure below is an example of the results we obtain from ELISA procedures to measure thermal degradation in milk immunoglobulins.

The figure below is a comparison of the 2D gels reported in the literature and the ones we have obtained in our laboratory. The procedure we are following is more direeted towards selection of MFGM only.

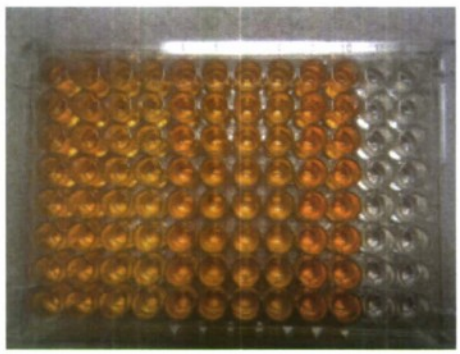


Two dimensional PAGE from MFGM. As reported by Fong et al.



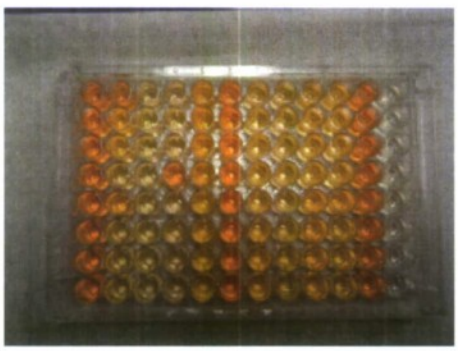
DPTC two dimensional gels of selected MFGM proteins

ELISA standards for determination of IgG in milk



Intact IgG react to give color. In this picture we have standards for calibration

Sample measurement of IgG in milk



Heated samples along high and low standards are measured in the ELISA plates

Task 6. Measure the immune-enhancing properties of MFGM fractions, before and after processing, by evaluating cytokine induction *via* RT-qPCR.
Part of this task was performed by our collaborators at biology (see report by Dr. Black).

This objective was based on the set up of the series of techniques needed to obtain a good two dimensional polyacrylamide gel electrophoretic (PAGE) pattern.

In summary, thanks to the work initiated by this grant we continue to have support for the application of a proteomic approach to the analysis of proteins interactions through processing.

Cal Poly Photosynthetic Wastewater Reclamation Project: Development of a Simple, Low-Cost Wastewater Treatment Process for Warm-Climate Communities and Military Bases

Project Investigator:

Tryg Lundquist
Civil and Environmental Engineering
California Polytechnic State University
San Luis Obispo, CA

Executive Summary

Algae-based wastewater reclamation processes are potentially important to the US Department of Defense in two main ways: as a cost-effective means of wastewater reclamation at military bases and as a source of biomass for biofuel production. The latter use of algae is the subject of research programs at DARPA and the Air Force. A key technical problem for both applications of algae technology is the high cost of harvesting the microscopic algal cells using mechanical or chemical means. Our present project has focused on developing methods to promote reliable bioflocculation of algae cells, which allows for their separation from the water by low-cost settling. In addition, we have monitored the effectiveness of algae in their wastewater treatment role. The experiments were performed in outdoor mixed tanks with municipal wastewater as the growth medium. Two different techniques to promote bioflocculation were investigated: high organic matter loading and inoculation with flocculating microorganisms. During the project, our understanding of how to use these methods has improved, and both have proven promising for algae harvesting, with a majority of final algae concentrations being <50 mg/L in the outdoor systems. This concentration is better than that in the effluents of most full-scale algae treatment systems in California, which typically have 80-120 mg/L of algae in their effluents. Wastewater treatment was also successful. Average organic pollutant removals ranged from 70-80% (as BOD_5), and average ammonium removal reached 64%. The ammonium removal was especially impressive since the systems were not being operated in a nutrient removal mode. Under continuing C3RP sponsorship, the bioflocculation and pollutant removal methods are being explored over a wider range of parameters. In addition to its research role, the pilot systems have been used in education and outreach. Visitors have included representatives of the PG&E Emerging Clean Technology Policy group, the Electric Power Research Institute, Carbon Capture Corporation, engineering firms, and numerous class field trips.

Acknowledgments

The following organizations and individuals have provided assistance to this project:

City of San Luis Obispo Wastewater Reclamation Facility Staff
Waste Connections, Inc.
Professor Yarrow Nelson
Professor Chris Kitts
Seppi Henemann
Stephen Barr
Kyle Poole
Christine Frost
Claire Balint

SECTION 1: BACKGROUND ON ALGAE BIOFUELS AND WASTEWATER TREATMENT

1.1 Problem Statement and Significance: Algae Biofuel

Algae are considered an attractive potential biofuel feedstock due to their high productivity compared to terrestrial crops and their often high oil content. For example, DARPA announced a $100 million solicitation entitled "BAA08-07, Biofuels–Cellulosic and Algal Feedstocks (Biofuels) Program" in November 2007. The relevant topic in the solicitation is "Processes for the Affordable Manufacture of Algal-Derived JP-8 Surrogate." Successful proposers are likely to be businesses that will be able to produce large volumes of algal triglycerides (ATGs) suitable for JP-8 jet fuel production. DARPA already has contracts in place for the production of JP-8 from ATGs, given their availability (Kirkpatrick 2007). Algae production on wastewater media is considered to be a competitive process by DARPA (Kirkpatrick, pers. com. 2007). In addition to the DARPA effort, the Air Force Office of Scientific Research is beginning to pursue basic research in the algae biofuel area. AFOSR held a research roadmap workshop in conjunction with the National Renewable Energy Laboratory in February 2008, which the PI attended. The main challenges are the production/harvesting of large quantities of suitable algae biomass and the extraction of the ATGs.

1.2 Problem Statement and Significance: Wastewater Treatment

Wastewater collection and treatment is one of the vital services that must be supplied by the US military at its bases and in its community development projects. In fiscal year 2003, the US military operated 700 overseas bases and another 6,000 within the US and its territories (DOD, 2003). The Army alone operates over 100 small wastewater treatment plants on bases (Scholze and Zaghloul, 2001). If similar to civilian communities, a substantial proportion of the public works budget of a military base is expended on the wastewater management enterprise. For development projects in countries where the electricity supply is unreliable and where operator training is minimal, conventional wastewater treatment facilities are likely to fail, despite considerable capital investments.

US civilian needs are also great. The American Society of Civil Engineers estimates that over the next 20 years communities in the US will need to invest $390 billion to renovate existing wastewater treatment facilities and to build an additional 2,000 treatment facilities (ASCE 2005). This level of investment has been called an "unprecedented financial problem" for utilities (WIN, 2000).

A subset of these facilities needing renovation includes some of the approximately 8,000 municipal wastewater pond systems in the US. The majority of these pond facilities are 10-30 years old, and most will need to be upgraded to improve effluent quality and to accommodate increased flows. Where land and climate allow, these needs can be met at low-cost with pond-based technologies such as the "high-rate pond" process under study in this C3RP project. High-rate ponds are shallow (1-3 ft) and mixed to promote algal growth.

Ponds are the third mostly commonly used treatment technology in the US. Mechanical technologies, such as activated sludge and trickling filters, are the most common means of treatment in the US. These technologies usually produce high quality effluent suitable for discharge to water bodies or for reuse. They are especially suited to cold climates (e.g., the northern US and Canada) and large cities where land is expensive. Unfortunately, mechanical technologies have serious drawbacks. Even considering the latest improvements in the technologies, they have problems such as the following (SBW Consulting, Inc. 2002; WEF, 1997; Green et al. 1995; Metcalf and Eddy, 2002; and USACE, 1999):

- High electrical energy consumption with concomitant greenhouse gas emissions (1,500–2,200 kWh per million gallons treated)
- High production of sludge (0.75–1.1 tons per million gallons treated)
- High capital costs due to intensive use of electric motors and concrete construction
- High operational costs due to the high number of operators and the high skill levels required of them
- Low reliability where the above conditions are not met

So called "natural" wastewater treatment technologies such as pond and wetland systems are an increasingly popular alternative to mechanical technologies because the natural systems use few mechanical parts, sludge is often disposed of onsite, operational requirements are minimal, and in some cases nutrients and energy can be recovered from the biomass produced during treatment. Natural treatment technologies use many of the same mechanisms of treatment that are used in mechanical systems. The main differences are that natural systems rely primarily on solar energy input rather than the electrical energy input of mechanical systems. Advantages of natural systems can include the following (Crites and Tchobanoglous, 1998; Downing et al. 2002):

- Production of biomass for bioenergy production
- Low total cost (a 4-fold advantage compared to the activated sludge process)
- Low energy demand, due in part to photosynthetic oxygen production
- Lower labor costs in terms of time required and worker skill required
- Little or no sludge handling and disposal cost in some cases
- Greater sustainability than mechanical treatment technologies
- Natural disinfection

Of course, natural systems have their drawbacks and perceived drawbacks as well, such as the following (USACE, 1999):

- Much greater land area requirement
- Effluent quality that does not meet modern regulatory requirements
- Perception of odor potential

Experience has shown that only poorly designed or grossly overloaded natural treatment systems emit odors (USEPA, 2007). The other two criticisms are valid. Wetland systems, though popular with the public, occupy much more land than equivalent pond systems due to their shallow water depth and long hydraulic residence times. Pond systems are a more compact technology, and high-rate pond systems have further decreased the areal requirements of ponds since their development during the 1950s- 1990s (Oswald, 1990; Green et al. 1995 & 1996). In the California climate, pond systems that incorporate high-rate ponds require 13–19 acres per million gallons per day (Fremming, Parson, and Pecchenino, Inc. 2005). The other valid criticism–poor effluent quality–is due mainly to microalgae cells suspended in the ponds effluent. In terms of dissolved pollutants, however, pond system effluent quality is often better than that of mechanical technologies (Green et al., 1996). Algal cells are difficult to remove due to their small size (1–30 μm diameter). Algal suspended solids cause the following problems:

- Exceedance of discharge limits
- Clogging of final effluent percolation basins
- Decreasing the effectiveness of disinfection processes
- Shortening facility lifetime due algal sludge in treated water storage ponds
- Promoting growth of cyanobacterial scums, which are an operational nuisance

The effluent quality limitations of ponds are due to suspended solids but also nutrients. However, high-rate ponds with sufficient inorganic carbon supplies are able to completely remove both soluble nitrogen and phosphorus (Nurdogan and Oswald 1995, Green et al. 1996, Woertz et al. 2008 in press).

Given the land efficiency and nutrient removal of high-rate ponds, ***high effluent suspended solids concentrations are the single most important obstacle to the widespread adoption of an otherwise highly advantageous wastewater treatment technology.*** In addition, ***the high cost of chemically-assisted harvesting of algae biomass is one of the most important obstacles to the production algae biomass for biofuels.***

Algal suspended solids can be removed from pond effluent by addition of advanced processes such as coagulation and flotation, but these more than double the capital and operations cost of advanced pond systems (Downing et al., 2002). Once harvested (by whatever means), algae can be anaerobically digested for methane production, as is done currently at full-scale at the City of Sunnyvale pond system (D. Eisenberg, EOA, Inc., pers. com.; BEI, Inc., and EOA, Inc. 1988). Production of liquid biofuel from algae is another route that is highly sought but not yet demonstrated, outside of the production of a few liters by proprietary processes (Bryan 2006).

1.3 Possible Means of Low-Cost Algae Harvesting

Evidence from several sources suggests that waste-grown algae cells can be caused to flocculate and rapidly settle without addition of chemicals (Garcia et al. 2006, Feffer 2007).

This type of flocculation is termed bioflocculation. Bioflocculation of *bacteria* is the basis of the highly-successful activated sludge process for wastewater treatment. If bioflocculation can be reliably induced in *algal-* or *algal-bacterial cultures*, then a similar demand for algae-based technology should follow. Such bioflocculation would also greatly decrease the cost of biofuel biomass production.

The cause and control of bioflocculation are topics requiring investigation. Statistical analysis of data from nine advanced pond or high-rate pond systems suggests that an unusual parameter strongly correlates with total suspended solids in the algae settling unit effluents (Lundquist, unpublished). The parameter, which apparently relates to bioflocculation, is the ratio of soluble biochemical oxygen demand concentration (SBOD) in the high-rate pond influent and total suspended solids concentration (TSS) in the high-rate pond effluent. The correlation is shown in **Figure 1-1**. The main goal of the present research is to test the hypothesis suggested by the graph: increased SBOD loading will improve settling below typical discharged limits of 40-45 mg/L. If reproducible at full-scale, this bioflocculation could save water districts chemical flocculant costs of $200-$400 per million gallons of wastewater treated.

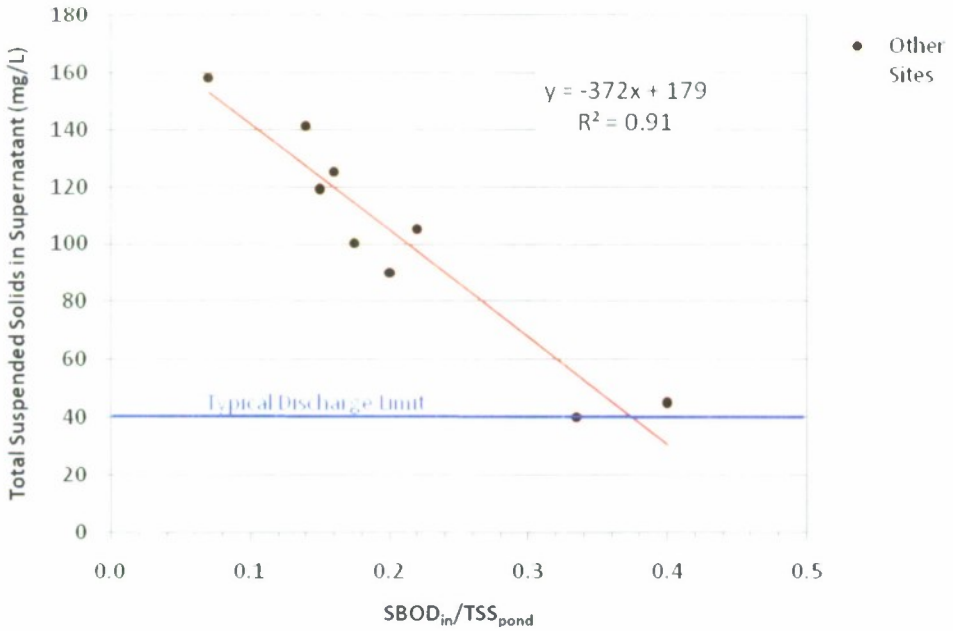


Figure 2-1. Correlation of the mean $SBOD_{in}/TSS_{pond}$ ratio to settling unit effluent TSS at nine different high-rate pond facilities operating during ten different periods. The correlation suggests that permit concentrations of 40-45 mg/L could be met if SBOD concentrations were increased, as described in the text. Pond locations were in California (Delhi, Hilmar, Richmond, Hollister, St. Helena), New Zealand (Ngatea, Hautapu), Spain (Barcelona), and South Africa (Grahamstown).

The hypothesized mechanism behind the relationship in **Figure 1-1** is that when sufficiently high, SBOD concentrations entering the HRP provide substrate (food) for bioflocculating bacteria such as are common in activated sludge facilities (e.g., *Zooglea*). The bioflocculating bacteria produce extra-cellular polysaccharide matrices that cause bacterial and algal cells to stick together forming flocs large enough to rapidly settle. Informally, SBOD is analogous to glue. The ratio of $SBOD_P/TSS_T$ indicates how much glue is present per particle needing glue. ***The common failure of high-rate pond systems to settle algae is probably due to excessive removal of SBOD in the primary ponds that are frequently used.*** The solution proposed is to replace primary ponds with small conventional primary clarifiers so that more SBOD remains to promote flocculation in the subsequent HRP. A secondary benefit of eliminating the primary pond is the decrease in overall land use. The typical excess oxygenation capacity of the HRP should be sufficient to satisfy the oxygen demand in the effluent of the small primary unit.

The main hypotheses listed in the 2006 C3RP proposal were the following:

Hypothesis 1: The total suspended solids concentration in settling tank effluents (TSS_S) is a function of the ratio of soluble biochemical oxygen demand concentration in the influent to the algae treatment pond ($SBOD_{in}$) and the total suspended solids concentration in the treatment pond (TSS_{pond}) but not a direct function of $SBOD_{in}$ or TSS_{pond} individually.

Hypothesis 2: TSS_S can be maintained consistently below the typically regulatory limit of 40 mg/L through manipulation of $SBOD_{in}/TSS_{pond}$.

Hypothesis 3: The settling tank effluent total BOD can be maintained consistently below the typically regulatory limit of 40 mg/L without increasing the treatment pond depth or hydraulic residence time outside the typical practical ranges of 0.3–1 m deep and 3–7 days, respectively.

1.4 Research Progress on Low-Cost Algae Harvesting under C3RP

During the Winter and Spring 2007 quarters, the PI designed the wastewater treatment-algae production pilot plant and ordered custom-made tanks, while the two MS students setup and learned analytical methods. During June-July, the PI and three students constructed the pilot plant at the City of San Luis Obispo Wastewater Reclamation Facility (WRF). The pilot plant consists of four 5-m^2 high-rate ponds and the systems to feed wastewater to the ponds and discharge the treated water back into the WRF system. This pilot facility is a primary tool in overall research program called the Cal Poly Photosynthetic Wastewater Reclamation Project, which also makes use of laboratory cultures and studies at full-scale wastewater treatment pond systems.

Pilot plant operation began in early August, and after a month of trouble-shooting and culture adaptation, routine sampling and experimentation began on Sept. 2, 2007. Samples were collected and analyzed for multiple water quality characteristics once per week.

The purpose of the study was to test hypotheses on the effect of SBOD/TSS on bioflocculation and, if possible, to demonstrate that low TSS could be achieved consistently. Several secondary hypotheses related to the rates of algal productivity and nutrient removal were also tested. Higher-than-normal SBOD/TSS ratios were provided through the use of primary effluent as the pond influent, but the original experimental plan had to be modified given the early results. Initially, the SBOD of WRF wastewater was not as high as expected on a consistent basis. As such, the plan to dilute the wastewater to achieve a range of SBOD/TSS ratios was not necessary. The natural variability of SBOD has provided the range of conditions.

In addition to the experimental use of primary effluent feed water, a second experiment was conducted on semi-continuous inoculation with flocculating microorganisms. The inoculum was return activated sludge from the WRF. Such organisms would be available at full-scale pond facilities if a small sidestream of wastewater was treated using the activated sludge process.

The following Materials and Methods section describes the pilot plant design, construction, and operation to test the research hypotheses.

SECTION 2: MATERIALS AND METHODS

Experiments were performed in four identical pilot-scale high-rate pond (HRP) reactors. The reactors were constructed during June-July 2007 at the City of San Luis Obispo Wastewater Reclamation Facility (WRF) in coastal central California (**Figure 2-1**).

The HRPs were used to study the effects of experimental and natural variables on water treatment performance and on the bioflocculation and sedimentation of algal-bacterial biomass. The HRPs were operated continuously during the experimental periods with continuous inflow and outflow. A process flow diagram of the pilot plant is shown in **Figure 2-2**. The four HRPs were named Northwest (NW), Northeast (NE), Southwest (SW), and Southeast (SE) based on their geographical position. Construction, operations and maintenance, as well as the procedures for experimentation, water quality analyses, and data analyses are described in this section.

2.1 Pilot Plant Construction and Layout

The four HRPs were placed in secondary containment (approximately 4,500 gallon capacity) (**Figure 2-3**) as a precaution to protect nearby storm drains. To prepare the site, a front loader tractor leveled the site, and nonwoven geotextile was laid down to protect the secondary containment liner from stones. The liner was 45-mil PVC donated by Waste Connections, Inc. The walls of the secondary containment were constructed of 2" x 12" lumber. Both ground wasps and gophers have been a problem at the site. Pest control contractors have been provided by the WRF.

Each high-rate pond was constructed from a rectangular fiberglass tank (10-ft x 5-ft x 2.5-ft deep) with a central baffle (King Starboard® marine plastic) and a custom paddle wheel (Starboard® blades bolted to an aerator paddle from Aquatic Ecosystems, Inc.). The paddle wheels were fixed to a stainless steel shaft powered by a 1/8-hp DC motor (Bodine, Inc.) with a power controller (Minarik Corp.) to provide gentle mixing to prevent algae from settling in the HRPs and to promote high algae productivity, as is done in full-scale HRPs. The paddle motor and controller were mounted to a plywood platform bolted to the lip of the tank (**Figure 2-4**).

For tank effluent, a 1.5-inch PVC stand-pipe was installed in each tank to maintain constant head. The standpipe, or effluent pipe, directed effluent flows through a PVC tank adapter, which reduced to a 1/4-inch hose barb fitted with 5/16-inch ID vinyl tubing. The vinyl tubing from all tanks directed the effluent flow to a sump in one corner of the containment area. The vinyl tubing clogged frequently during the first months of operation and was replaced by 3/4-inch PVC pipe (**Figure 2-5**). A sump pump, with a float switch, directed water from the sump back to the main flow of the WRF (**Figure 2-6**).

The pilot system received primary effluent from one of the WRF primary clarifiers. A ¼-hp submersible pump placed in the weir of the primary clarifier (**Figure 2-7**) delivered the water to the pilot system constant-head feed tank. The pump directed the primary effluent through 3/4-inch polyethylene irrigation tubing, which was threaded through 1-inch secondary

Figure 2-1. Aerial view of the City of San Luis Obispo Wastewater Reclamation Facility showing the location of the pilot plant and the primary clarifier which provided wastewater feed to the pilot plant. Source: Google Earth.

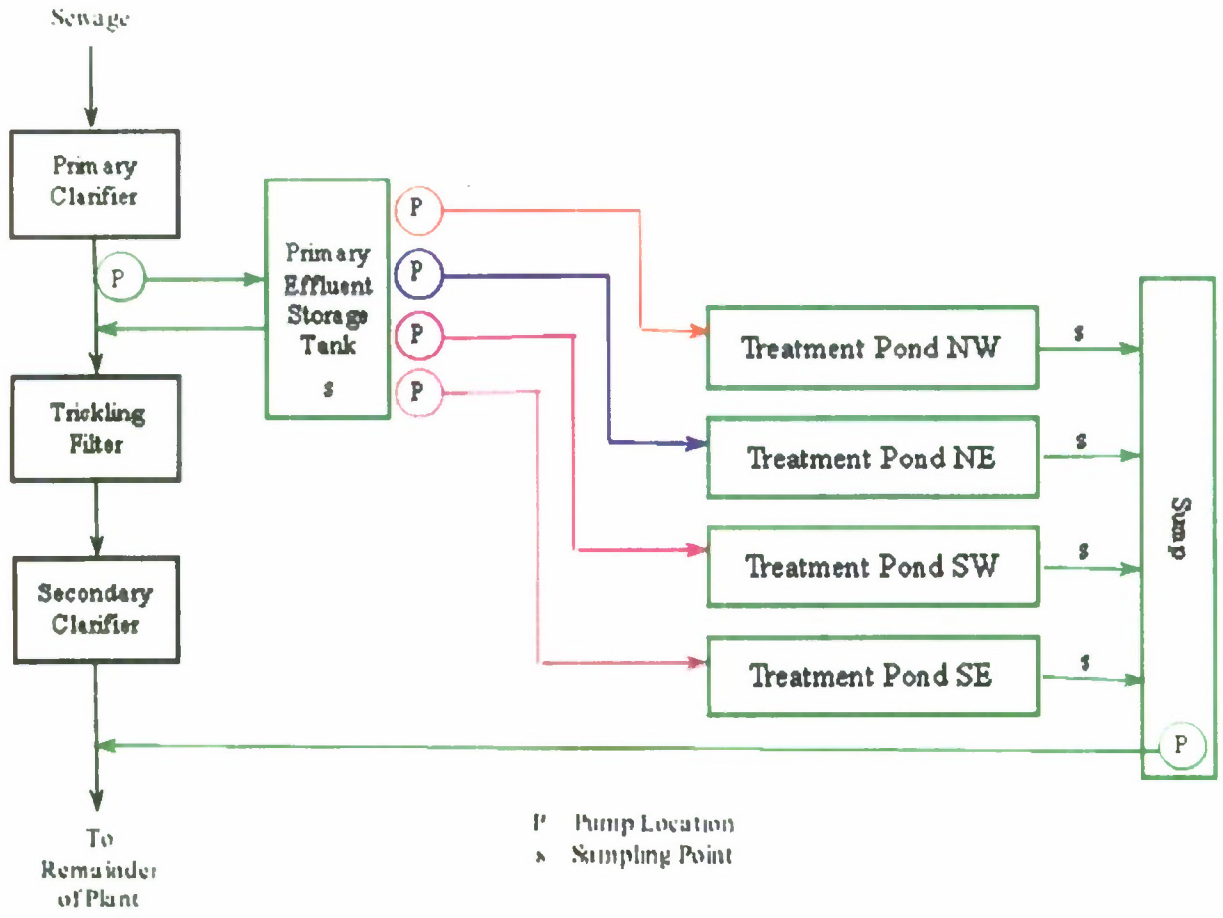


Figure 2-2. Schematic of the pilot plant and its connection to the SLO Wastewater Reclamation Facility.

Figure 2-3. Placement of HRPs in secondary containment using knuckle boom provided by the SLO WRF.

Figure 2-4. A high-rate pond tank during test filling, showing the central baffle, paddle wheel, and motor.

Figure 2-5. HRP effluent tank adaptor, tubing, and piping.

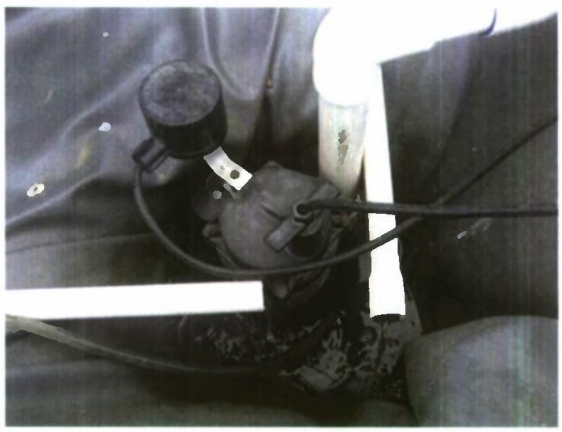


Figure 2-6. Effluent piping and sump pump with a float switch.

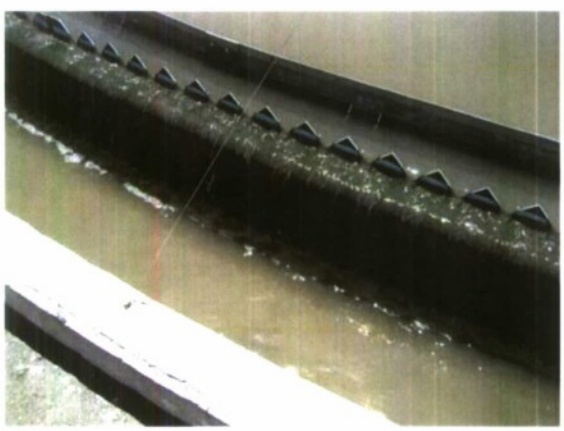


Figure 2-7. Primary clarifier weir from where pilot plant influent was pumped.

Figure 2-8. Influent tubing from the clarifier crossing the secondary containment wall.

Figure 2-9. Influent constant-head feed tank with inflow and outflow piping back to the clarifier.

Figure 2-10. Top view of influent feed tank with mostly submerged standpipe shown.

containment tubing (**Figure 2-8**), to the constant-head feed tank (20 gallons) located in the center of the pilot facility (**Figure 2-9**).

In order to prevent changes in primary effluent quality during transport to the ponds, it was necessary to limit the hydraulic residence time (HRT) of the tubing and feed tank. A short HRT (~10 min.) was achieved by pumping a high flow to the feed tank and allowing excess overflow to return by gravity to the clarifier weir via a 3/4-inch PVC effluent standpipe in the feed tank (**Figure 2-10**). The primary effluent in the feed tank continuously flowed through a 1/2-inch PVC manifold to four peristaltic pumps that individually fed the HRPs. The flow influent flow rates provided a residence time of 3 days during most of the 2007 work. The pond water depth was 2.2 ft (68 cm), and the volume was 580 gallons (2,200 L) for each pond.

2.2 Pilot Plant Operations and Maintenance

2.2.1 Velocity Gradient Calculations

Gentle mixing is well-known to promote flocculation in water treatment. The flow mixing of HRPs is an example of such gentle mixing. However, the power required for flow mixing is a cubic function of flow velocity. With increasing mixing power, there is a trade-off between improved flocculation and power use. Pilot systems can be easily over-mixed, which may cause better flocculation than would be achieved at full-scale. It was important that the low-speed mixing used in full-scale HRPs was reproduced in the pilot ponds despite the huge differences in pond size.

The root-mean-square (RMS) velocity gradient (G) is a common parameter used to characterize mixing intensity and to design mixing systems (MWH 2005). The velocity gradient is a function of the power dissipated per unit volume, and it can be used to scale back and forth between pilot-scale and full-scale systems. The RMS velocity gradient is calculated with the following equation.

$$\overline{G} = \sqrt{\frac{P}{\mu V}} \qquad \text{(Equation 2-1)}$$

where $\overline{G}$ = RMS velocity gradient, s^{-1}

μ = dynamic viscosity of water, $N \cdot s/m^2$

P = power of mixing input to vessel, J/s or W

V = volume of mixing vessel, m^3

For systems with rotational mixing devices such as paddle wheels, the power dissipation can be estimated with the following equation (MWH 2005):

$$P = \frac{C_D A_p \rho v_r^3}{2} \qquad \text{(Equation 2-2)}$$

where C_D = drag coefficient on paddle, unitless

A_p = projected area of paddle, m^2

p = fluid density, kg/m^3

v_r = velocity of paddle relative to fluid, m/s

A typical velocity gradient for a full-scale high-rate pond facility is 0.29 s^{-1}. This velocity gradient was reproduced in the pilot ponds using the above equations as a guide. To have G = 0.29 s^{-1} in the pilot ponds, a paddle wheel rotational speed of 2.0 rpm was needed. The details of this calculation will be shown in the MS thesis stemming from this project (Frost 2008, in preparation).

2.2.2 Hydraulic Retention Time

In the California climate, high-rate ponds treating sewage have design hydraulic residence times (HRTs) of 3-10 days, depending on season and treatment objectives. A 3-d HRT was used in the pilot ponds during most of 2007 in order to test the most land-efficient HRT and to provide a high SBOD loading.

2.2.3 Start-Up Procedure

On August 9, 2007, prior to the beginning of experimentation, all but one HRP (SE) was filled with half reclaimed water and half primary effluent and inoculated with 200 mL of algae. The algae inoculum consisted of samples collected from several wastewater treatment pond systems in California, as well as a sample from a local creek. The prominent algae genera in the inoculum were *Actinastrum, Scenedesmus, Chlorella, Spirogyra, Nitzschia, Micractinium, Golenkinia, Chlorococcum, Closterium, Euglena*, and two unidentified species (Feffer 2007). After inoculation, the HRPs were operated with a HRT of 2.4 days for three weeks and a paddle wheel speed of 2 rpm while the pilot facility construction was completed.

After three weeks the ponds were green with phytoplankton algae, and the tanks were intermixed to provide homogeneity as the first experimental period began. On August 25, 2007, the HRPs were intermixed using the SE pond as a mixing and transfer tank. Then, all the ponds were brought to the same depth (2.2 ft) by filling with primary effluent. Finally, each pond was given an initial inoculation of 1.5 gallons of return activated sludge (RAS) from the WRF nitrifying aeration basins. The inoculation with RAS ensured that the ponds had typical flocculation-promoting organisms such as ciliates, rotifers, and many others.

2.2.4 Routine Monitoring and Maintenance

The HRP operation required monitoring of two operator-controlled variables—flow rate and paddle wheel rotational speed. These were monitored on a nearly daily basis, and adjusted as needed to maintain steady operation (**Table 2-1**). Influent flow to each pond was measured by

collecting the influent flow in a 500-mL graduated cylinder over the course of 60 sec. Paddle velocity was timed as seconds per one rotation.

Influent and effluent tubing and piping were inspected frequently for leaks or clogging and repaired or cleared when necessary. In warmer months, filamentous algal-bacterial growths formed on the upper few centimeters of the submerged tank walls. These were brushed off as needed. Removal of the filamentous growth was always performed more than three days prior to a sampling date to minimize their influence on TSS measurements.

Less frequent maintenance tasks included cleaning and lubrication or replacement of poorly performing pumps and motors.

Table 2-1. Operational variables during Aug. 26-Dec. 12, 2007 for the high-rate pond pilot systems.

Variable	NW	SW	NE	SE
Flow (mL/min)*	504	518	494	502
Std. Deviation	75.1	42.0	88.0	40.6
No. Measurements	79	79	79	79
Paddle Speed (rpm)	2.1	2.0	2.1	2.0
Std. Deviation	0.73	0.42	0.61	0.29
No. Measurements	107	107	107	107

* From Sep. 24 - Dec. 12, 2007

2.3 Experimental Procedures

The purpose of the following experiments was to test the ponds for their response to input variables in terms of bioflocculation and sedimentation of biomass and in terms of treatment performance. This section outlines the procedures followed for gathering baseline data, as well as the experimental procedures for Experiments I and II.

2.3.1 Baseline Water Quality Data

Knowledge of the characteristics of the WRF primary and secondary effluents was needed to finalize the pond and experimental designs. (Secondary effluent was the planned dilution water.) For this purpose, BOD and solids were determined approximately bi-monthly from February-August 2007, when the pilot system was completed. Monthly WRF influent water quality data for 2001-2006 was also obtained in order to indicate the variability of total BOD, of which soluble BOD is a subset. These data will be discussed in Frost 2008, in preparation.

2.3.2 Initial Operations and Determination of Variability

From August 25-September 23, 2007, the four ponds were operated identically to determine the natural variability among the cultures and over time. This information is useful in estimating the uncertainty of the later experimental results, as will be examine in Frost 2008. With a 2.4-day HRT (flow rate = 630 mL/min/pond), treatment performance was characterized

in terms of BOD removal, ammonia removal, biomass production, and biomass settleability. Temperature and pH of the pilot plant influent and in each HRP was measured. Flow rate, paddle wheel speed, and weather data were collected daily, while water quality parameters were generally determined once per week.

2.3.3 Experiment I

Experiment I had two purposes related to bioflocculation, settling, and treatment performance: (1) to gather more data at high SBOD/TSS ratios of 0.35 and greater and (2) to determine the effects, if any, of daily inoculation of the ponds with floc-promoting organisms. All four ponds received primary effluent, which was higher strength than the wastewater normally fed to HRPs, so the first purpose was achieved in all four ponds. For the second purpose, two ponds were inoculated daily, and the other two acted as controls.

For the inoculated ponds, 1 gallon/day of return activated sludge (RAS) from the WRF nitrifying aeration basin was added to the SW and NW ponds. The NE and SE ponds received no RAS. The NW and SW served as duplicate experimental inoculated ponds, while the NE and SE were considered duplicate control ponds. The RAS was obtained by repeatedly dipping a 1/2-gallon sampling bucket into the RAS feedwell of the WRF activated sludge systems.

All four HRPs were operated with a 3-day residence time (510 mL/min/pond) and a paddle wheel speed of 2 rpm. Flow rates and paddle wheel speeds were adjusted daily as necessary to maintain target values.

Experiment I began on September 24, 2007 and operated continuously until December 12, 2007. Floc quality was determined by field and laboratory settling tests and microscopic observations. Treatment performance was evaluated by BOD removal, ammonia removal, settling efficiency, and soluble nutrient removal. Temperature, pH, dissolved oxygen, and in-pond settled sludge were measured periodically. Additional water quality data for the mixed liquor and RAS from the activated sludge basin were obtained from WRF staff.

2.3.4 Experiment II

Experiment II is ongoing and is similar to Experiment I except that the HRTs in the four ponds have been increased to 5 days. A longer HRT should produce more flocculation and increased DO concentrations. The increased flocculation would occur due to an older average age of the cells (similar to what occurs in the activated sludge process) and due to the longer average time that the cells are undergoing flocculation mixing. The higher DO would occur due to the decreased BOD mass loading on the ponds. Experiment II began in January 2008 and is only brief preliminary results are presented in this report.

2.4 Water Quality Sampling and Analysis

The following sections outline the various water quality analytical techniques used in the evaluation of treatment performance for the HRPs. The constituents analyzed are listed in **Table 2-2**.

Table 2-2. Experiment I Constituents Determined and Sampling Schedule.

Parameter	Locations Sampled	Measurement Frequency
TSS/VSS	INF, HRPs	Weekly*
TDS/VDS	INF, HRPs	Once per Experiment
BOD	INF/HRPs	Weekly*
sBOD	INF/HRPs	Weekly*
nBOD/cBOD	INF	Weekly*
Ammonia	INF/HRPs	Weekly*
Soluble nutrients	INF/HRPs	Weekly*
Total phosphorus	HRPs	Once per Experiment
Alkalinity	HRPs	Once per Experiment
Dissolved Oxygen	HRPs	Occasional
Diel DO readings	HRPs	Once per Experiment
pH	HRPs	Weekly*
Temperature	HRPs	Weekly*
Field settling tests	HRPs	Daily
Laboratory settling tests	HRPs	Weekly*
Algae ID & Enumeration	HRPs	Once per Experiment

* Measurement performed weekly, with exceptions. Power outages and mechanical failures in the month of November resulted in the cancellation of several samplings.

Grab sampling in HDPE containers was used exclusively. Each HRP was sampled by dipping a 1-gallon container 1-cm below the pond surface to simulate flow over the effluent standpipe. The samples were collected at the center of the HRP channel upstream of the influent point.

The HRP influent was collected at a sampling port in the piping manifold that fed the peristaltic pumps. Prior to sampling, the sampling port valve was opened half-way to gently purge the manifold. Then the sample was collected. Sample containers were rinsed once with the water being sampled prior to use and were then stored on ice for no more than 30 minutes prior to arriving to the laboratory. After water quality analysis, each sample container was brushed and then rinsed with deionized water prior to reuse.

2.4.1 Solids

Total suspended solids (TSS), volatile suspended solids (VSS), total dissolved solids (TDS), total solids (TS), and volatile solids (VS) were performed following APHA Methods 2540 B, C, D, and E. Filters were pre-washed and ashed. A TSS standard was frequently used for quality control. The filters used for solids tests were Fisher Scientific G4 glass fiber filters with a nominal pore diameter of 1.2 µm (equivalent to Whatman GF/C). The filtrate from the G4 filtrations was used for the dissolved solids tests.

2.4.2 Optical Absorbance

Optical absorbance was explored as an alternative to TSS tests, but the correlation between absorbance and TSS was not strong, probably due to the presence of flocs. The spectrometer (Bausch and Lomb Spectronic 21) was turned on 20 minutes before each use to allow the bulb to warm-up prior to analyzing samples. The spectrometer was set to read absorbance at 750 nm and medium sensitivity, as described in Feffer 2007. After blanking the spectrometer with deionized water in a cleaned 5-mL glass cuvette (1-cm diameter), 5 mL of gently-mixed sample was placed in a glass cuvette and inverted five times prior to reading its absorbance. Due to variability in the absorbance caused by flocs, the highest observed value was always recorded for consistency.

2.4.3 Biochemical Oxygen Demand

Total, soluble, carbonaceous, and nitrogenous biochemical oxygen demands were determined per APHA 5210 B. Soluble BOD was determined using the filtrate from 1.2-µm Fisher G4 glass fiber filters. Hach nitrification inhibitor (Formula 2533™) was used to determine carbonaceous BOD on both filtered and unfiltered samples. Nitrogenous BOD was determined as the difference between total and carbonaceous BOD, and soluble nitrogenous BOD was determined as the difference between soluble and soluble carbonaceous BOD. Standards and blanks were analyzed in each batch of samples.

2.4.4 Colored Organic Matter Fluorescence

Raw fluorescence was measured with a Turner Designs Trilogy Laboratory Fluorometer with a colored organic matter module in an attempt to develop a correlation between raw fluorescence and soluble BOD. Raw fluorescence measurements were abandoned after early results indicated that no significant correlation existed between soluble BOD and raw fluorescence.

2.4.5 Settling Tests

Settleability was determined in 100-mL beakers, 500-mL beakers, and 1-L Imhoff cones, for comparison. (Concurrent laboratory studies required the use of 100-mL beakers for settling.) Supernatant samples were collected after 2 hours of settling and after 24 hours. The initial sample and supernatants were analyzed for total suspended solids concentration and optical absorbance. The Imhoff cones were less efficient settling containers compared to the beakers, but the Imhoff cones were selected for routine use since they are the standard device for determining settleable solids in wastewater laboratories (APHA Method 2540 A).

For the settling test, the sample bottle was gently inverted three times, and then the sample was poured into the settling device. At 2 hours and 24 hours later, 100 mL of supernatant was carefully pipetted from 1-cm below the water surface to avoid re-suspension of settled flocs or entrainment of any floating material.

After 24 hours of settling, the final volume of settled biomass in the Imhoff cones was recorded. This measurement estimated the biomass volumes produced and also may be useful

for calculating a parameter similar to the sludge volume index, which is used in activated sludge operations.

Informal 20 min. settling tests were also performed in the field using 100-mL glass beakers. The general speed of settling, floc characteristics, and supernatant color and clarity were recorded.

2.4.6 In-pond Settled Biomass

The mass of algal-bacterial flocs that settled in the ponds was determined with the aid of a Sludge Judge® II. The sludge judge was used to grab cross-section samples of the tank to determine the depth of sludge settled at the bottom of the HRPs. Samples were collected at several locations in each pond: in the channel opposite to the paddle wheel, along the baffle wall, and in the corners. Sludge depth was recorded as an average of three samples taken at three different locations within each section (middle, baffle, and corner) sampled. Total and volatile solids of the collected sludge was determined per APHA.

2.4.7 Ammonium

Ammonium was determined using the Ammonia-Selective Electrode Method (APHA 4500-NH_3 D.). After bringing samples to room temperature, Orion Ammonia pH-adjusting ISA (#1310-73-2) was used to raise the pH of the sample or standard prior to measurement. Calibration curves were developed using 1 mg/L, 10 mg/L, 50 mg/L, and sometimes 100 mg/L ammonia as N standards made from 1000 mg/L stock solution. The ammonia probe was rinsed between readings using deionized water. Ammonium was measured in samples of the HRP influent and HRP effluents for most sample dates.

2.4.8 Soluble Nutrients

Nitrate, nitrite, and phosphate were determined using ion chromatography, as described in Feffer 2007.

2.4.9 Total Phosphorus

Total phosphorus was determined using the Ascorbic Acid APHA Method 4500 E. Total phosphorus measurements were performed occasionally during each experiment.

2.4.10 Alkalinity

In order to characterize the influent and effluent waters, alkalinity was measured using APHA Method 2320 B. Titration Method.

2.4.11 Algae Identification and Enumeration

Algae identification and enumeration was performed at least once per experiment. Algae identification was performed by taking photomicrographs of flocs and the dominate types of algae. Algae were identified to the genus level using information in Standard Methods and other identification materials. The photomicrographs were taken at 100x, 400x, and 1000x total magnifications using an Olympus CX41 optical microscope with phase contrast coupled

with an Infinity 2 digital camera and Infinity Analyze software. The scale in the Infinity Analyze software was calibrated using the 50-µm squares of a Brightline hemacytometer.

Algae enumeration was performed by placing approximately 100 µL of vigorously mixed sample onto a Brightline hemacytometer with a hemacytometer cover slip. The sample was pipetted into the capillary notch of the hemacytometer and allowed to enter the counting chamber. The counts were done at 400x total magnification on the smallest squares in the counting chamber (0.0025 mm^2). The algae quantities and genera observed were recorded until at least 200 cells were counted and identified. If the minimum quantity of 200 algae was not accomplished after counting all 400 small squares, then algae were identified and counted in the larger 0.04 mm^2 outer squares. The concentration of the most abundant identified genera were calculated by dividing the number of each species counted by the area of the hemacytometer used for enumeration (the total area of all the small and large squares counted), multiplying that value by the total area of the counting chamber and then dividing that product by the volume of sample in the counting chamber (100 µL). Algae genera and density results will be discussed in Frost 2008, in preparation.

2.4.12 Temperature and pH

Temperature and pH were measured in the field using a portable Oakton Acorn® Ion 6 meter. The Oakton meter was calibrated periodically according to the manufacturer's instructions. Measurements were made directly in the feed tank and HRPs. Measurements were made to determine if the HRPs ever exceeded a pH of 9. At such high pH values, autoflocculation may occur, which would confound the bioflocculation results.

2.4.13 Dissolved Oxygen

Dissolved oxygen (DO) was measured in the field using a portable YSI 58 DO meter. The YSI meter was calibrated according to the manufacturer's instructions prior to each use. Measurements were taken by rinsing the DO electrode with deionized water and placing the electrode directly in to the HRPs and gently moving the electrode back and forth in the water. The electrode was rinsed with deionized water and air-dried prior to placing back in its storage bottle.

2.4.14 Weather Data

Daily solar insolation, precipitation, evapotranspiration, average weekly air temperature, and average wind speed were obtained from the California Irrigation Management Information System through the Department of Water Resources Office of Water Use Efficiency online database. The data was obtained from San Luis Obispo Station No. 52 located on the California Polytechnic State University campus (35°18'22"N, 120°39'37"W), which is approximately 3 miles from the pilot plant.

2.4.15 Statistical Analysis

All water quality analytical data, field observations, weather, and maintenance data were uploaded to statistical software (JMPIN, Version 4.04, SAS Institute) for multivariate analysis and modeling, to be described in Frost 2008 (MS thesis in preparation).

SECTION 3: RESULTS

The pilot plant work has generated data and observations on algae bioflocculation, floc settling efficiency, and treatment performance. The basic variables under our control were the hydraulic residence time (two levels tested and a third underway), depth (fixed at 68 cm or 2.2 ft), and paddle wheel speed (fixed at 2 rpm). The environmental variables, over which we did not have control, were the influent water quality and the weather. The following sections describe the results during the period of experimental operation—September 2-December 12, 2007. Operation restarted in January 2008, but most of the continuing work will be described in subsequent reports to the C3RP program.

3.1 Environmental Conditions

Temperature and insolation (solar energy flux) have important influences on treatment in algae systems. Warmer temperatures speed treatment, and high insolation promotes increased algae productivity and photosynthetic oxygen production. Although mean daily air temperatures were variable from day to day, the average temperature did not change much over the course of the experiments. September temperatures averaged about 17°C, while December temperatures averaged about 13°C (**Figure 3-1**). Insolation ranged from 36-264 W/m^2 (**Figure 3-2**). Net evaporation was estimated as precipitation minus reference evapotranspiration. Evaporation off the surface of the ponds amounted to about 2% of the wastewater inflow (**Figure 3-3**). These meteorological data were downloaded from a weather station on the Cal Poly campus, as described in the Materials and Methods section.

The primary clarifier effluent delivered to the high-rate ponds generally increased in strength during the course of the experiments. Total BOD_5 ranged from 140-450 mg/L (**Figure 3-4**). Typically, a majority of the total BOD_5 was comprised of soluble BOD_5. The carbonaceous BOD_5 concentration was quite steady at 130-175 mg/L (**Figure 3-5**).

Total suspended solids in the influent to the high-rate ponds generally ranged from 80-130 mg/L, with 80-90% comprised of volatile suspended solids (VSS) (**Figure 3-6**). These values are typical of primary clarifier effluents at municipal wastewater treatment plants. The extent to which these volatile solids are degraded in the high-rate ponds or are transmitted to the effluent will be estimated in Frost 2008, in preparation. This estimate will allow the newly grown algal-bacterial component of the effluent suspended solids to be differentiated from suspended solids of sewage origin and RAS origin, allowing a better estimate of the algae productivity of the ponds.

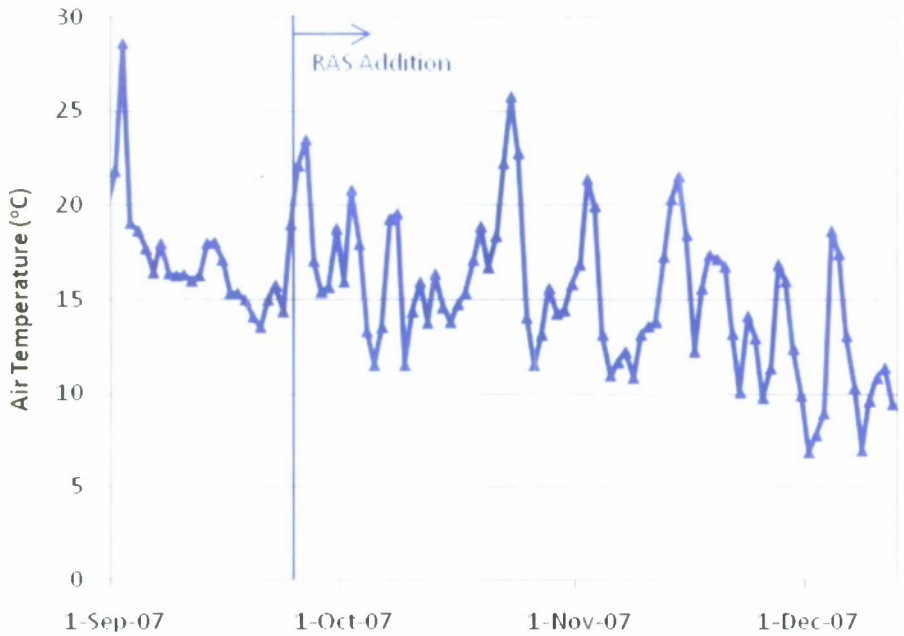


Figure 3-1. Mean daily air temperature in San Luis Obispo. The vertical line shows when the return activated sludge inoculation (RAS) experiment began.

Figure 3-2. Mean daily solar flux in San Luis Obispo. The vertical line shows when the return activated sludge inoculation (RAS) experiment began.

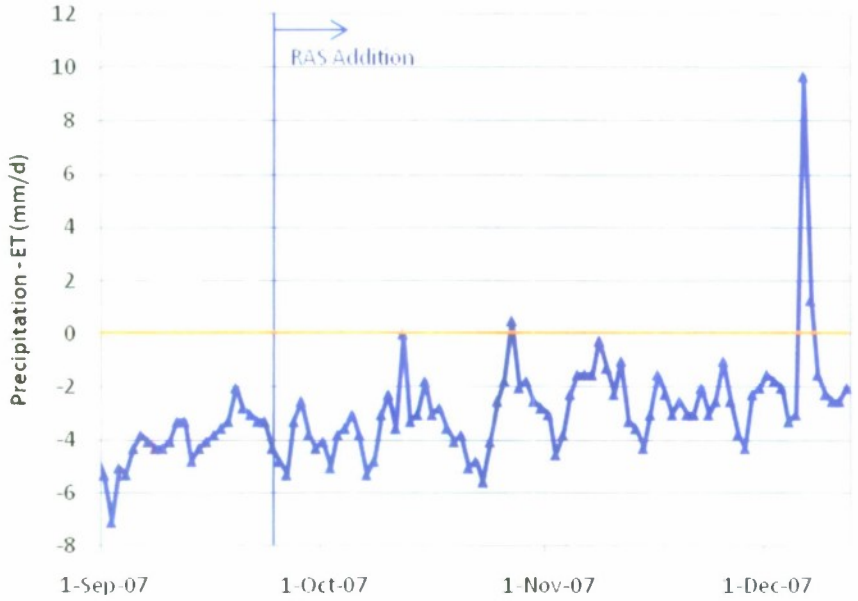


Figure 3-3. Approximate net evaporation (precipitation – evapotranspiration) in San Luis Obispo. The ponds lost an average of about 3 mm of water per day, which is equivalent to 15 L/pond/day or 2% of the inflow.

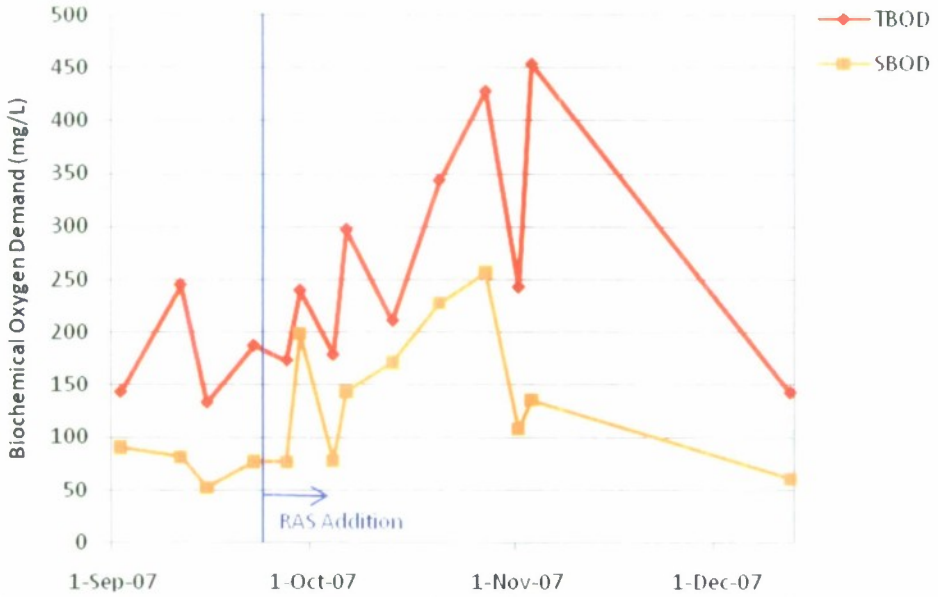


Figure 3-4. Total and soluble 5-day biochemical oxygen demand in the influent to the high-rate ponds. Soluble was separated from particulate with standard suspended solids filters (Fisher G4).

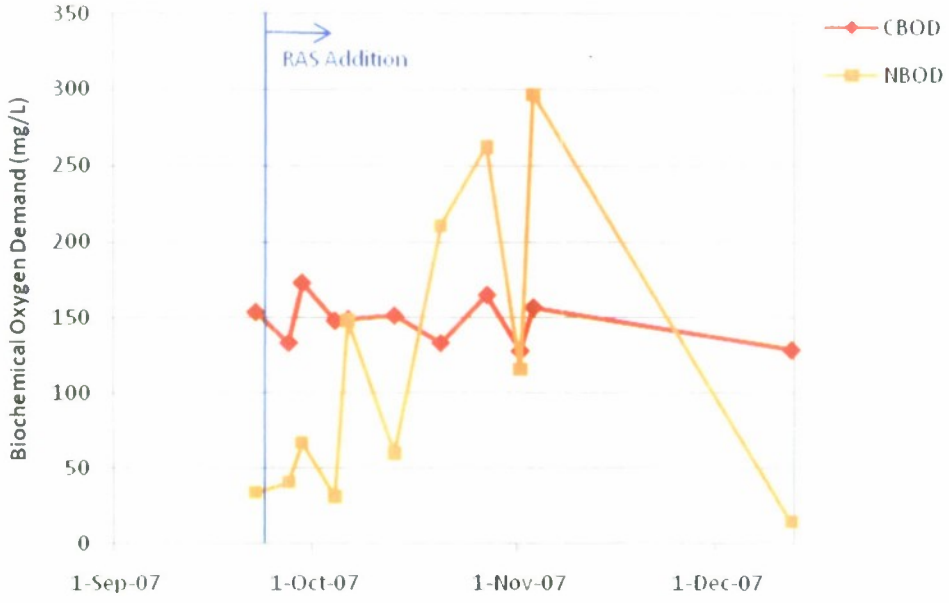


Figure 3-5. Carbonaceous and nitrogenous BOD5 in the influent to the pilot high-rate ponds.

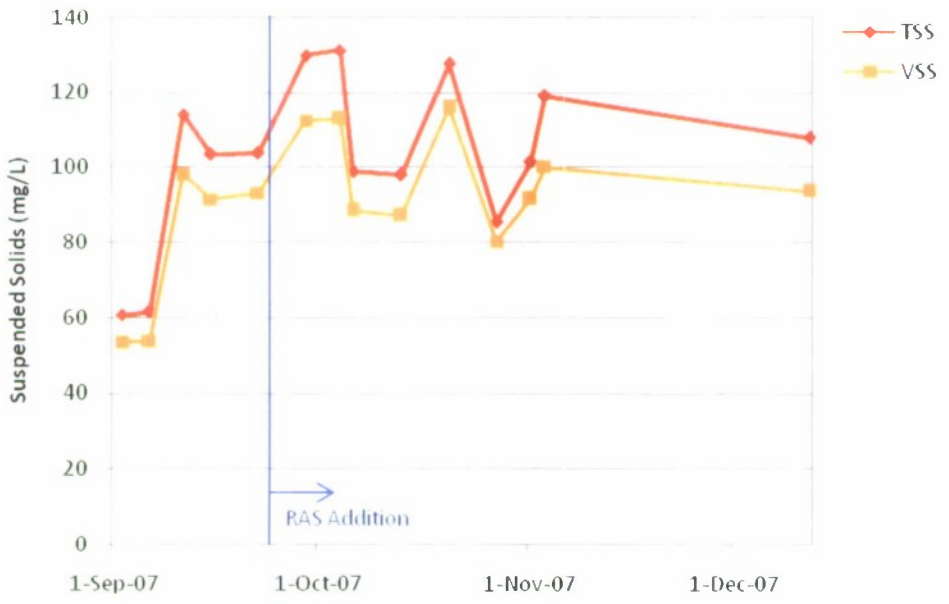


Figure 3-6. Total and volatile suspended solids concentrations in the influent to the pilot high-rate ponds.

3.2 Treatment Performance

We monitored total BOD and ammonium concentrations in addition to the suspended solids and soluble BOD, which were the main focus of our research (**Figure 3-7, Table 3-1**). Mean SBOD removals ranged from 69-80% in the four HRPs, with average effluent concentrations of 25-38 mg/L. These levels of SBOD removal are similar to conventional facultative wastewater treatment ponds, but the HRPs achieved the removal with a 5-fold shorter hydraulic residence time. For the HRPs, total BOD removal would not be as complete as the SBOD removal due to oxygen demand of the effluent algal cells. However, this BOD from residual algae will be minor when more complete algae settling is achieved, as has been the case during the early 2008 experiment.

Ammonium removal in HRPs is accomplished by two main mechanisms: assimilation into algal cells and volatilization out of the pond's surface. Nitrification of ammonium to nitrate is another mechanism, although it is rarely an important factor in ponds. Volatilization is also minor when water pH is less than 9, as was the case in our pilot HRPs. Thus, assimilation into algal biomass was the main removal mechanism. Ammonium removal ranged from 40-64% without consistent correspondence to the control or inoculums ponds. Although not complete, this substantial ammonium conversion to algal biomass is beneficial considering that ammonium removal by conventional mechanical technologies is particularly energy-intensive. We expect complete ammonium removal to be accomplished during the high productivity months of summer. Operation with longer residence times and/or with carbon dioxide addition are other promising methods to achieve complete nitrogen and phosphorus removal (Feffer 2007).

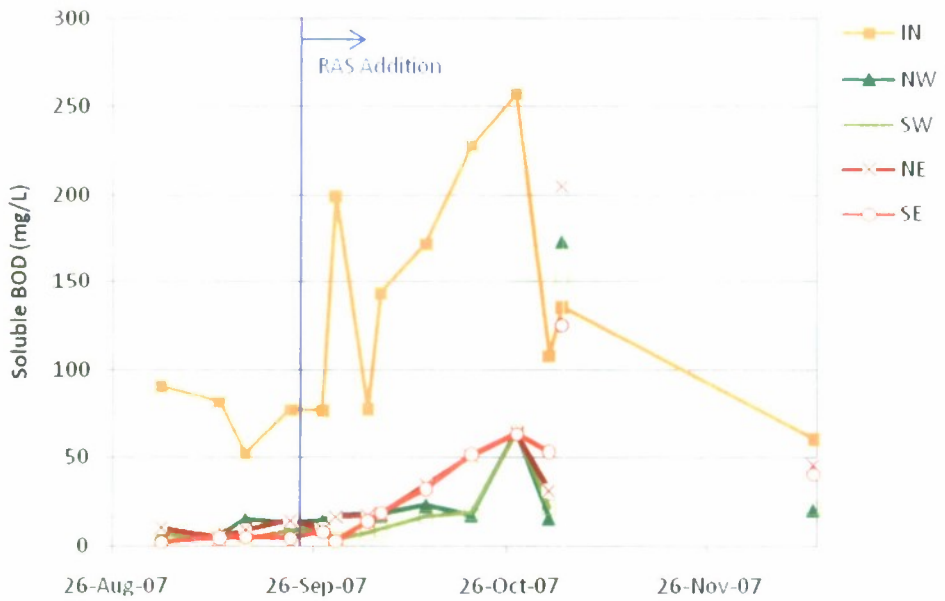


Figure 3-7. Soluble BOD in the influent and effluent of the pilot high-rate ponds.

Table 3-1. Average concentrations of water quality constituents during Aug. 26-Dec. 12, 2007 in the high-rate pond pilot systems.

Constituent	Influent	NW	SW	NE	SE
TSS (mg/L)	103	229	208	136	165
Std. Error	5.9	16.6	10.6	16.3	19.8
No. Samples	14	15	15	15	15
VSS (mg/L)	91	201	186	123	147
Std. Error	5.0	14.9	9.7	13.9	17.3
No. Samples	14	15	15	15	15
$TBOD_5$ (mg/L)	244	ND	ND	ND	ND
Std. Error	27.4	ND	ND	ND	ND
No. Samples	14	ND	ND	ND	ND
$SBOD_5$ (mg/L)	126	30	25	38	31
Std. Error	17.5	11.6	10.6	13.7	9.3
No. Samples	14	14	14	14	14
% Removal	NA	75%	80%	69%	75%
$CBOD_5$ (mg/L)*	147	91	75	73	87
Std. Error	4.6	11.9	3.9	9.4	1.8
No. Samples	11	3	3	3	3
$NBOD_5$ (mg/L)*	116	176	113	104	96
Std. Error	30.1	73.2	64.1	51.9	38.6
No. Samples	11	3	3	3	3
Tot. Amm. (mg/L)	38	23	14	23	17
Std. Error	1.7	1.3	1.0	1.1	1.9
No. Samples	12	12	12	12	12
% Removal	NA	40%	64%	40%	55%

* Pond results are in terms of soluble $CBOD_5$ and soluble $NBOD_5$.

ND = Not Determined. NA = Not Applicable.

3.3 Biomass Productivity and Wastewater Treatment

During September, the control pond effluent VSS concentrations increased substantially above the influent VSS due to growth of algae and bacteria. However, as insolation and temperature declined into winter, the VSS increase in the control ponds also declined, as would be expected. The inoculated ponds experienced a similar decline, but their starting VSS concentrations were much higher than the control ponds. **Figure 3-8** shows this trend for the North ponds. The South pond data are similar, and are shown as an overlay in **Figure 3-9**.

Beyond examining effluent VSS concentrations, we can discuss two types of "productivity." For ponds, productivity is expressed as grams produced per square meter per day ($g/m^2/d$) indicating the dependence of productivity on insolation. The first type of productivity is "solids production" which is based only on the effluent mass flow rate of VSS and includes residual

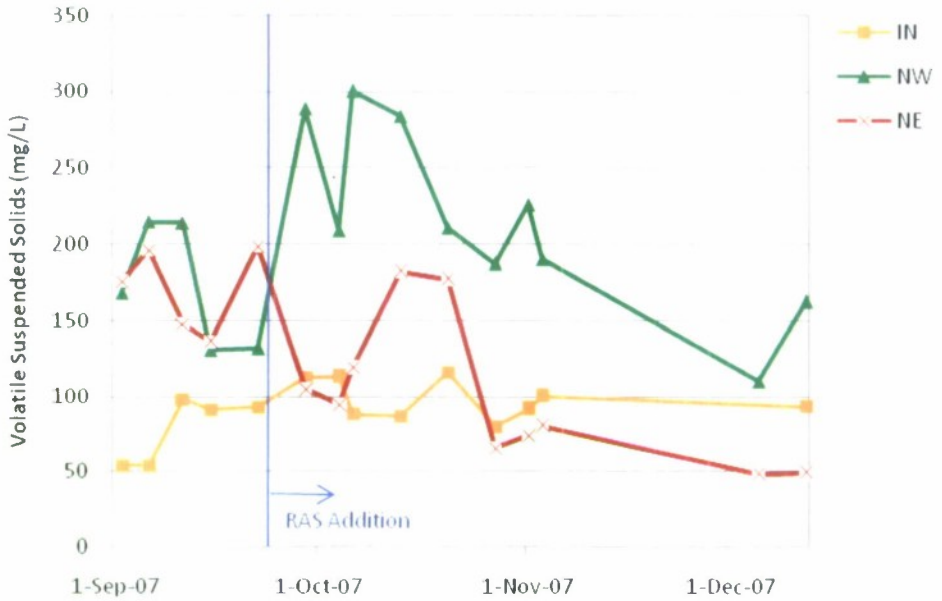


Figure 3-8. Volatile suspended solids concentrations in the influent to the ponds and in the effluent of the NW (inoculated) and NE (control) ponds.

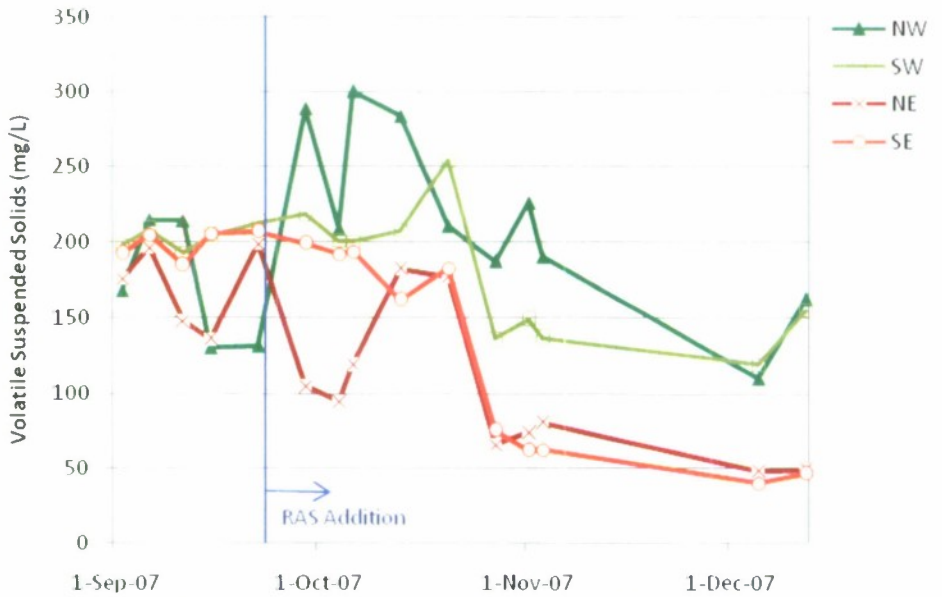


Figure 3-9. Volatile suspended solids concentrations in the effluent of all four ponds. NW and SW were the inoculated ponds, and NE and SE were the control ponds.

VSS of sewage and RAS origin. Solids production is important information for biomass management, including biofuel and fertilizer uses of the solids. The second type is "net cell productivity," which is based on the net gain or loss of VSS. In other words, the net change in VSS concentration equals the following:

(Effluent VSS conc.) – (Influent VSS conc.) – (VSS increase due to RAS addition)

Net cell productivity is important in terms of autotrophic algae growth, with its concomitant oxygen production, nutrient uptake, and CO_2 fixation. The heterotrophic growth of bacteria also contributes to cell productivity. The heterotrophic growth in the ponds can be estimated independently of the autotrophic growth by multiplying the SBOD removal by a typical bacterial yield (0.35-0.6 g VS produced per g BOD removed).

During the initial period of operation, when daily inoculation was not used in any of the ponds, the solids production was similar in all ponds (**Figure 3-10, left side**). However, the north ponds produced about 6 g/m^2/d less (15% less) than the south ponds. This consistent difference may have been due the different location of the paddle wheels in the north and south ponds. In the north ponds only, the paddle wheels shaded the return channel, which could have decreased algae growth. During the daily inoculation experiment, the north-south pond production difference was not consistent (**Figure 3-10, right side**).

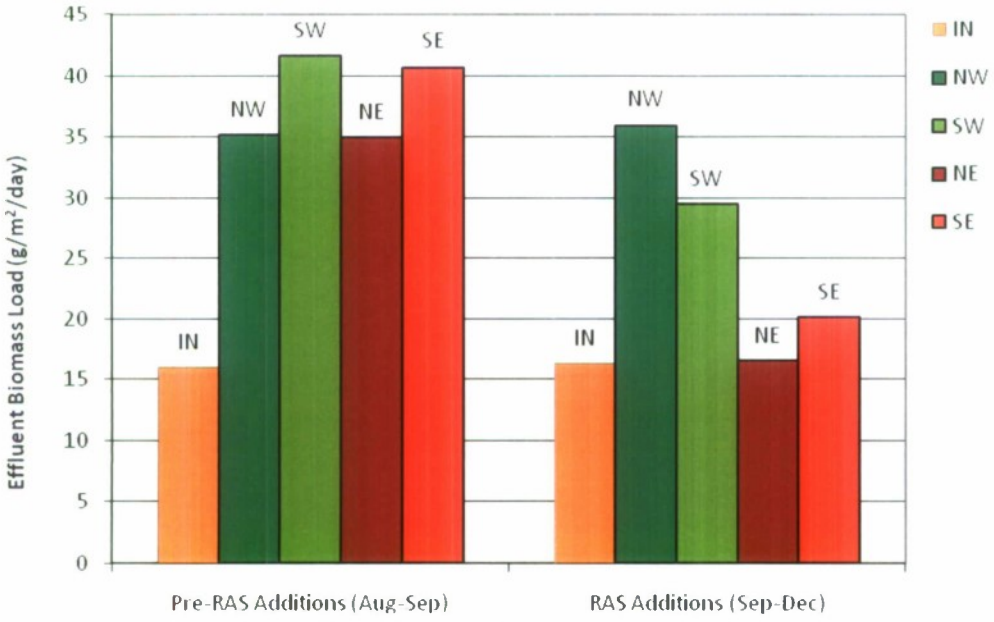


Figure 3-10. Solids production during the initial period prior to daily RAS addition (left) and during the RAS addition period (right). Solids production is expressed per unit area of pond and includes any residual VSS from the pond influent and RAS additions.

During the inoculation experiment, an average of 31 mg/L VSS (standard deviation=3.4 mg/L) was added to the inoculated ponds (west) with the RAS inoculations. This added concentration was 10-30% of the effluent VSS, depending on season. As described above, for the cell productivity calculations, this 31 mg/L VSS and the influent VSS concentration were subtracted from the effluent VSS concentration.

A stand-out result from the cell productivity calculations was the low net cell productivity of the control ponds (not graphed). The NE pond productivity was 0.23 $g/m^2/d$, while the SE productivity was 3.9 $g/m^2/d$. Despite these low calculated values, it was obvious from the dark green color of these ponds that algae growth was profuse, however, on a net VSS basis, productivity was low. It is evident that influent VSS were degraded in the pond and replaced by a nearly equal concentration of algal VSS.

Cell productivity in the inoculated ponds was much higher than the control ponds and more typical of high-rate ponds in the San Luis Obispo climate. The NW pond productivity was 15.6 $g/m^2/d$, and the SW pond productivity was 9.1 $g/m^2/d$. Since influent VSS no doubt degraded in the inoculated ponds as it did in the control ponds, actual algae productivity was somewhat higher than indicated by the values based strictly on net increase in VSS concentration.

The large difference in cell productivity in the control and inoculated ponds was not expected. A possible explanation is that the inoculated RAS organisms reproduced quickly thereby quickly increasing the bacterial VSS concentration. However, this explanation is not supported by the SBOD removal data.

(As mentioned above, heterotrophic cell growth is dependent on removal of substrate, SBOD in this case. Total BOD removal is not as relevant as SBOD removal because TBOD includes the oxygen demand of particulates, which are accounted for in the VSS concentrations.)

The inoculated ponds removed only slightly more SBOD than the control ponds (110 mg/L vs. 100 mg/L) and so should have grown only slightly more bacteria. Assuming a typical heterotrophic cell yield of 0.5 g VSS grown/g BOD removed and ignoring cell decay, the additional SBOD removal by the inoculated ponds would have amounted to only 5 mg/L more bacterial growth or 0.8 $g/m^2/d$. Subtracting this bacterial growth, the remaining difference in productivity between the control and inoculated ponds averaged 9.5 $g/m^2/d$.

This 9.5 $g/m^2/d$ productivity difference is substantial considering that typical algae productivity in high-rate ponds is 10-15 $g/m^2/d$ during winter in California. Either RAS inoculation was a remarkable stimulant to algae productivity or other less straightforward processes were at work. One possible explanation is that the control ponds were somehow inhibited. Their average productivity of 2.1 $g/m^2/d$ was much lower than normal, whereas the inoculated ponds had a net productivity of 11.3 $g/m^2/d$, which is typical. The control and inoculation ponds were

carefully operated to have nearly identical conditions, including fully intermixed initial cultures, so the cause of any inhibition is unknown.

Another explanation (though an unlikely one) for this apparent difference in productivity is that slowly biodegradable compounds were adsorbed to filterable solids (flocs) in the inoculated ponds and not to filterable solids (mostly colloids) in the control ponds. This differential absorption would have the effect of increasing the suspended solids concentration in the inoculated ponds. If the adsorbed compounds were slowly- or non-biodegradable, they would not be measured as BOD removal. This set of circumstances would have resulted in removal of soluble BOD that was neither accounted for in the BOD analysis nor in the estimate of bacterial growth due to BOD removal.

At this point, the reason for the large difference in productivity between the control and inoculated ponds remains unknown. The current experiment involves switching which ponds are inoculated. Results from this new experiment may shed some light on the productivity question.

3.4 Bioflocculation and Settling

A major goal of this project was to develop a method to cause algae to bioflocculate reliably, thereby allowing a clarifier to produce low TSS effluent (<40 mg/L). The main hypothesis to be tested was that high concentrations of soluble organic matter and/or low pond TSS concentrations promote low TSS clarifier effluents. In other words, we wanted to test whether high SBOD/TSS ratios promoted flocculation and settling.

The denominator of this ratio, TSS concentration, could only be influenced, not controlled. Beyond our control were major factors in TSS such as insolation, temperature, and zooplankton grazing. The variables with the greatest influence on TSS that *were* in our control were depth and hydraulic residence time. These two variables were held fixed throughout the experiment.

The original proposal envisioned that the influent wastewater SBOD concentration would be fairly constant and could be diluted with WRF secondary effluent to achieve several experimental levels of SBOD concentration. Controlling SBOD concentration would provide substantial control over the SBOD/TSS ratios, which were to be spread over a range that was higher than that studied previously (**Figure 1-1**). However, during the first five weeks of operation, we found that influent SBOD concentrations were lower than expected and that SBOD/TSS ratios (at 0.2-0.5) were only slightly higher than the previously tested ratios. For this reason, dilution would not have been helpful. In the following months, the SBOD/TSS ratio did increase several-fold, but we decided not to initiate dilution. Since influent SBOD was not steady, steady dilutions could not be achieved. Instead, we monitored pond performance given the naturally variable conditions. Because the influent was primary clarifier effluent,

rather than the partially-oxidized facultative pond effluent typically fed to high-rate ponds, the conditions in our ponds spanned the desired testing range.

In our experiments, a standard laboratory settling device, the Imhoff cone, was used to compare the settleability of the pond biomass (**Figure 3-11**). Imhoff supernatant TSS was measured at 2 h and 24 h after settling was started. For simplicity, only the 24-h data are discussed here. Both the control and inoculated ponds had similar 24-h supernatant TSS concentrations, but the inoculated ponds values averaged less than the target concentration of 40 mg/L (NW: 33 mg/L & SW: 36 mg/L), while the control ponds averaged greater than 40 mg/L (NE: 47 mg/L & SE: 41 mg/L) (**Figures 3-12 & 3-13**).

A more pronounced difference between the control and inoculated ponds was settling efficiency (% TSS removal). For all ponds, settling efficiency was higher when initial TSS concentrations were higher. Since the inoculated ponds maintained a higher initial TSS concentration than the control ponds through much of the study, the inoculated ponds had greater removal efficiencies (average removal of 84% compared to 70% in the control ponds).

The apparent relationship of supernatant TSS and the SBOD/TSS ratio (**Figure 1-1**) was tested in the current study by extending the SBOD/TSS range of previous experience. The highest ratio tested previously was 0.4, whereas in the current study, average ratios ranged from 0.3 to 1.5, with nearly 4 as a peak in one sample. The SBOD/TSS ratio generally increased through late October and then decreased (**Figure 3-14**) due in part to a similar pattern in influent SBOD concentrations (**Figure 3-7**).

The settling results are described in two periods: prior to the inoculation experiment and during the inoculation experiment. For the pre-inoculation period, the new SBOD/TSS data are shown in **Figure 3-15** as averages from each pond. These new data are plotted with the historical data seen previously in **Figure 1-1**. Although the SBOD/TSS ratios only extended up to 0.42 in this period, the new data filled in a region for which there were only two previous data points.

Figure 3-15 shows a strong linear correlation between the SBOD/TSS ratio and supernatant TSS (R^2=0.92). Most importantly, the new supernatant TSS concentrations were all <40 mg/L, a typical effluent TSS discharge limit. The new data suggests that if the SBOD/TSS ratio is maintained between 0.33-0.42, supernatant TSS values will usually be <40 mg/L. This more extensive data set will be important to the existing full-scale high-rate pond facilities in California, which are typically operating at a low SBOD/TSS (<0.3) and often discharging 80-120 mg/L TSS.

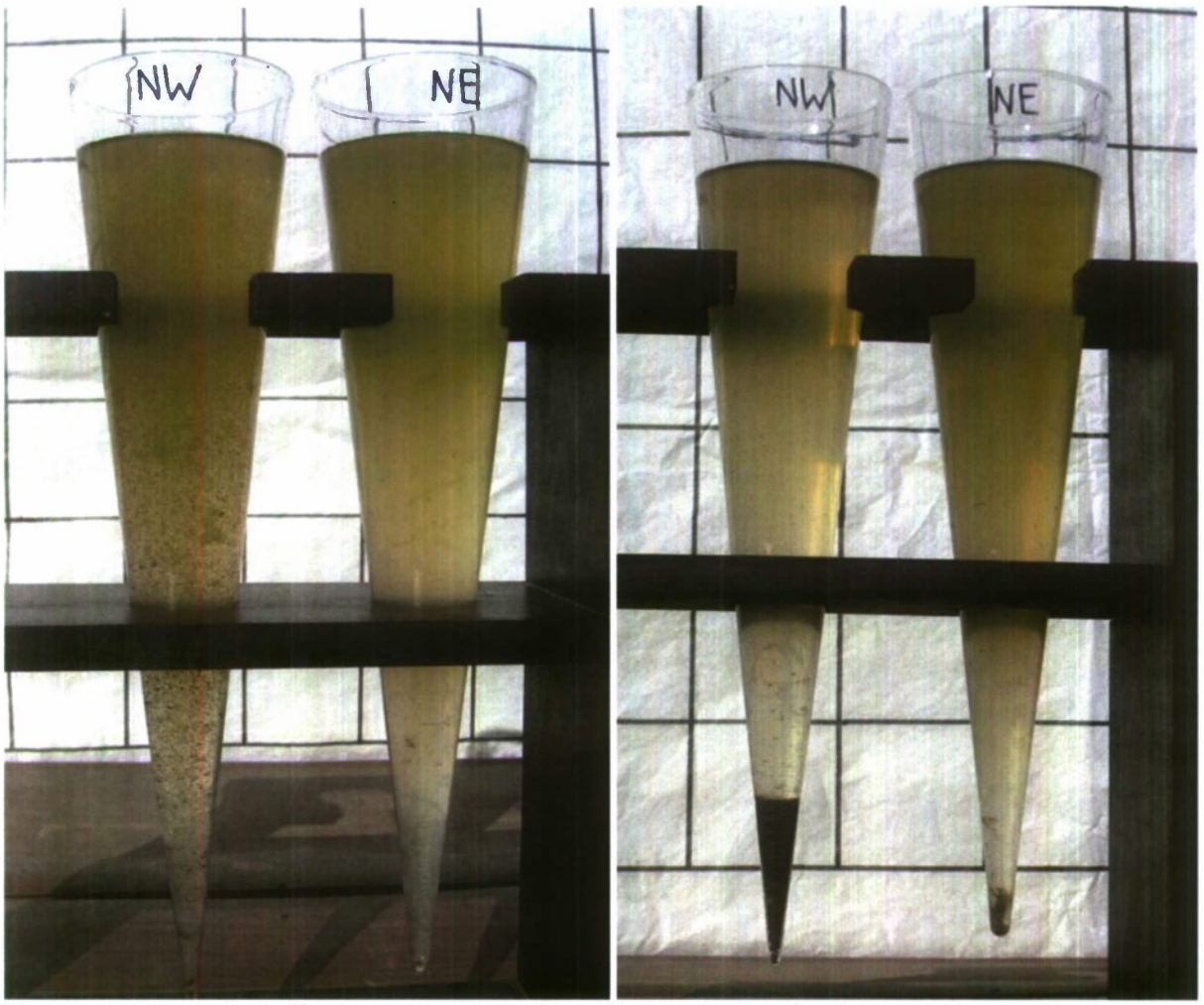


Figure 3-11. Typical Imhoff cone settling results during Experiment I (Left: Time=0 hrs, Right: Time=24 hrs). Both NW and NE ponds received primary effluent with elevated SBOD concentrations. In addition, the NW pond received daily RAS inoculation. The inoculated ponds displayed better flocculation and more settled biomass than the control ponds. On the right, the better supernatant clarity of the inoculated pond is apparent. (The settling test pictured was conducted over December 12-13, 2007. Such tests were conducted indoors with only the photographs taken outdoors.)

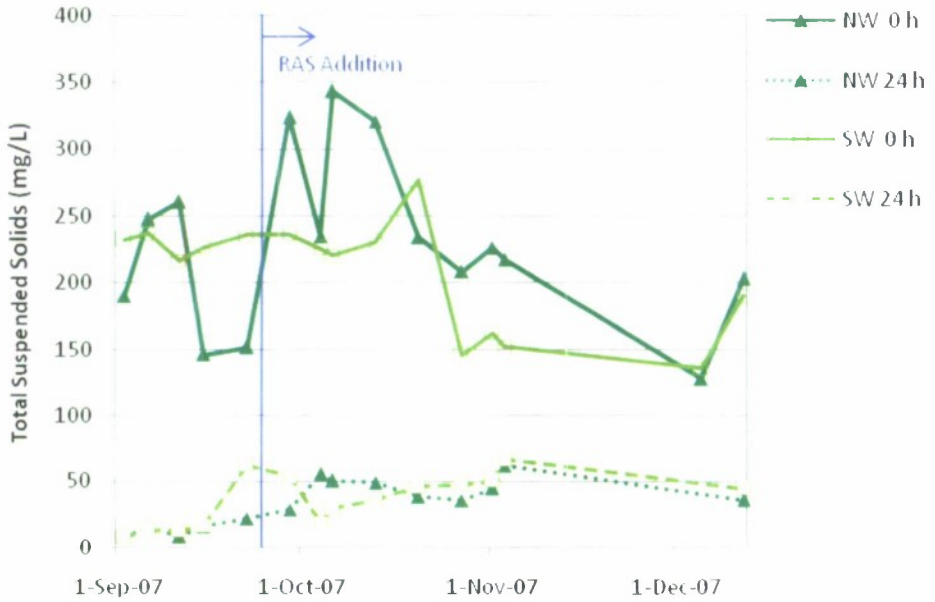


Figure 3-12. Total suspended solids concentrations initially and after 24-hours of Imhoff cone settling for the inoculated ponds.

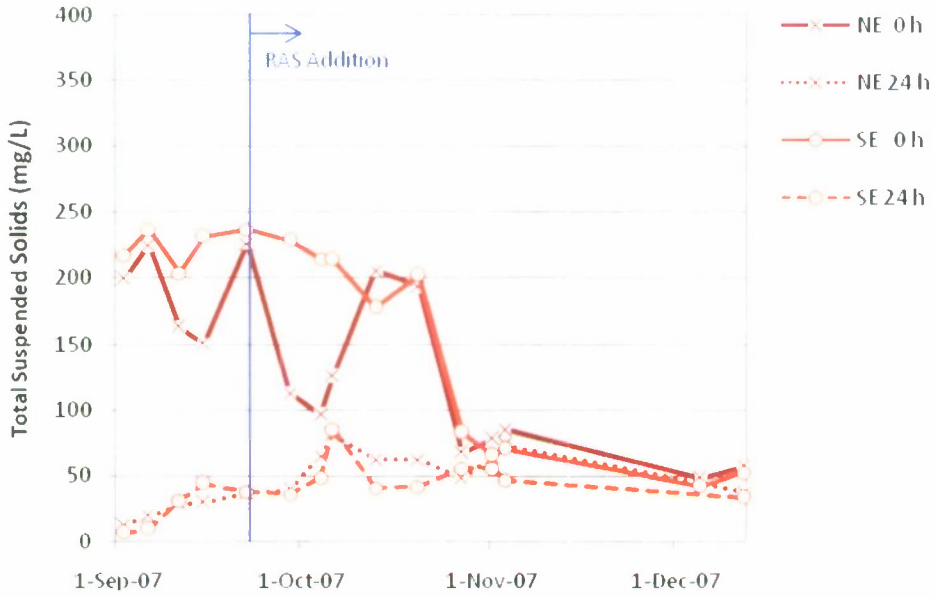


Figure 3-13. Total suspended solids concentrations initially and after 24-hours of Imhoff cone settling for the control ponds.

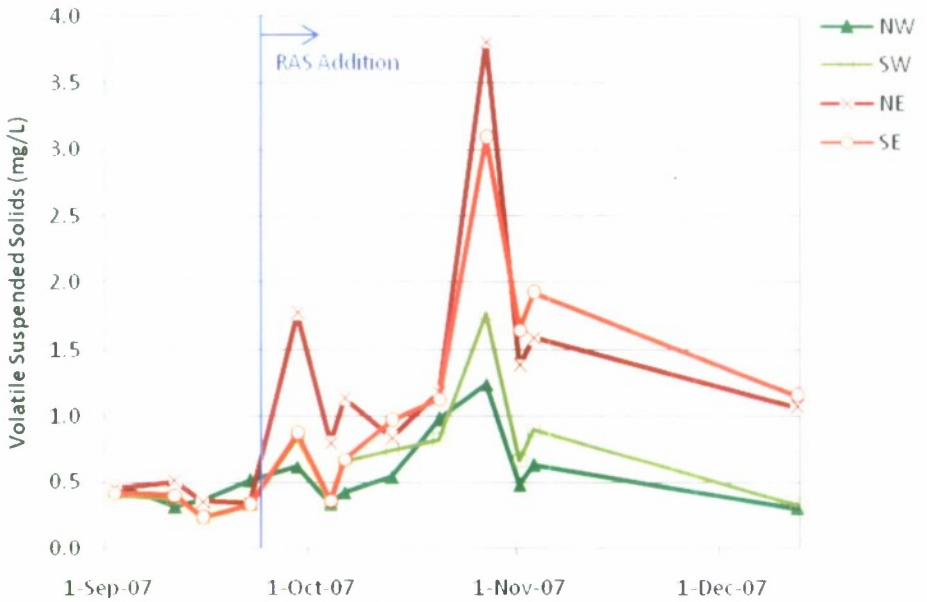


Figure 3-14. The time series of soluble BOD/TSS ratios for the inoculated and control ponds.

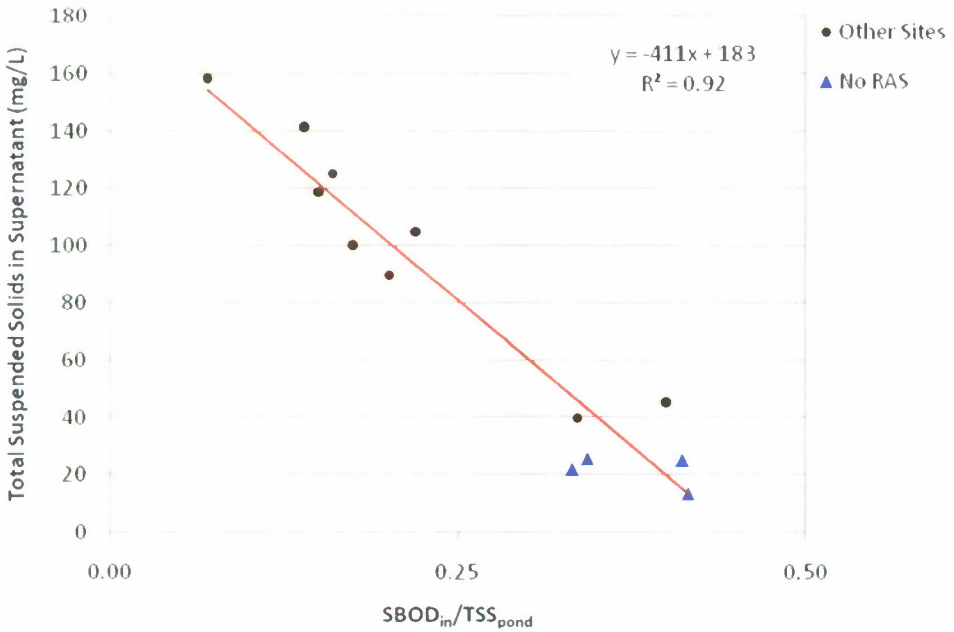


Figure 3-15. Correlation of supernatant TSS concentration with the soluble BOD/TSS ratio for the San Luis Obispo pilot high-rate ponds and for high-rate ponds at other sites. Ratios of 0.3-0.4 led to supernatant TSS concentrations less than the 40 mg/L target. Above a ratio of 0.5, supernatant TSS increased as seen in Figure 3-16. See Figure 1-1 for information on the historical data.

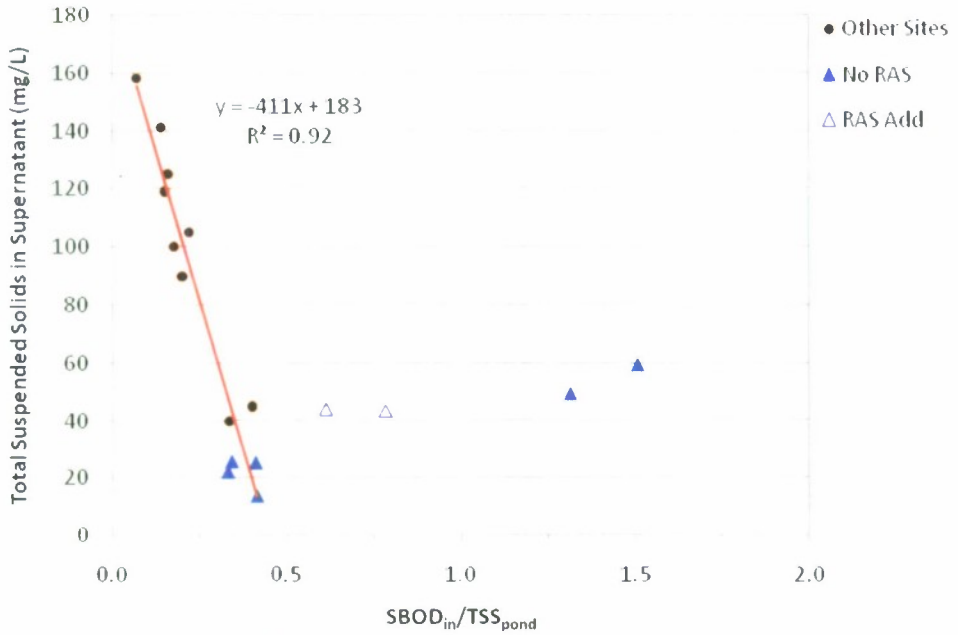


Figure 3-16. Correlation of supernatant TSS concentration with the soluble BOD/TSS ratio for the San Luis Obispo pilot high-rate ponds and for high-rate ponds at other sites, including the data corresponding to ratios >0.5. The linear relationship between supernatant TSS and the SBOD/TSS ratio breaks-down for high ratios. See Figure 1-1 for information on the historical data.

With further increasing SBOD/TSS ratios, we expected that at some point supernatant TSS concentrations would have to again increase, probably due to over-growth of colloidal bacteria. The linear relationship found for the low ratios was expected to transition into a somewhat parabolic relationship at the high SBOD loading rates. This expectation was partially fulfilled. During the daily inoculation experiment, the average SBOD/TSS ratios were far higher than the previous period, ranging from 0.6-1.5. At these elevated ratios, the SBOD/TSS-to-Supernatant TSS correlation was absent (**Figure 3-16**). Instead, the supernatant TSS concentrations increased slightly with higher SBOD/TSS ratios, similar to the expectations. However, at the high ratios, the supernatant still had a green tint suggesting that algae still dominated the TSS rather than the expected bacterial domination at high SBOD loading. The average supernatant TSS concentrations during this second experimental period were above the target at 43-59 mg/L. However, RAS addition appeared to provide a benefit of 10-20 mg/L lower supernatant TSS concentrations than the control ponds, which also had ratios well above 0.4.

3.5 Floc Appearance: Macroscopic and Microscopic

In searching for the mechanisms of improved flocculation performance, we have conducted extensive microscopic and field observation of the cultures. During the 2007 work, large dark-green flocs (3-5 mm) were seen in all ponds, in addition to the colloidal algae cells that colored

the water green. However, in the control ponds, the flocs were either nearly absent or present at very low densities (only several flocs per 10 mL). In the inoculated ponds, the flocs were present at a much higher density (dozens per 10 mL). This condition was probably the cause of the lower 24-hr supernatant TSS concentrations for the inoculated ponds (**Figure 3-11**).

Occasionally, for a several days at a time, the inoculated ponds developed copious tan-colored flocs (1-3-mm) numbering in the thousands per 10 mL. The floc density was so great, that no open water could be seen between the flocs when looking through a 100-mL beaker. During these times, settling was rapid and especially effective, with the supernatant containing a residual of only 10-20 g/L TSS.

Under the microscope, both the green and tan flocs were seen to consist of a matrix of algal and bacterial cells that appears similar to those of the activated sludge process. However, the tan flocs appears to have a core of algal cells surrounded by layer of bacterial cells. Understanding and controlling the formation of these two types of flocs will be helpful to the reliable operation of high-rate ponds. The copious tan flocs probably resulted in lower final TSS concentrations because the small flocs have a large surface area for adsorption of unflocculated cells. The layered appearance of the tan floc may be due to the timing of changes in wastewater quality (rate of bacterial food provision), the type of bacteria present, etc. A close analysis of the daily field observations in conjunction with weather and influent water quality data may suggest some mechanisms. This analysis will be described in Frost 2008, in preparation.

During 2008 when flocculation has been consistently good, we have made micrographs of the flocs of control and inoculated ponds (**Figure 3-17, top**). Although the alga *Actinastrum* sp. dominated both cultures, the inoculated ponds flocculated much more completely. The control ponds, especially NE, contained mostly individual cells or cells loosely clustered (**Figure 3-17, bottom**). The inoculated ponds, however, contained large algal-bacterial flocs and high concentrations of floc-promoting activated sludge organisms, such as stalked ciliates, paramecia, and rotifers (**Figure 3-17, top**). The good flocculation performance with inoculation conforms to the basic ecological principles of culture enrichment. What remains to be seen is the reliability of the flocculation over the seasons and over periods with widely varying wastewater quality. Among the factors needing to be optimized are the relative mass of the inoculated organisms needed and the timing/frequency of inoculations. These questions are being pursued in the continuing 2008 work.

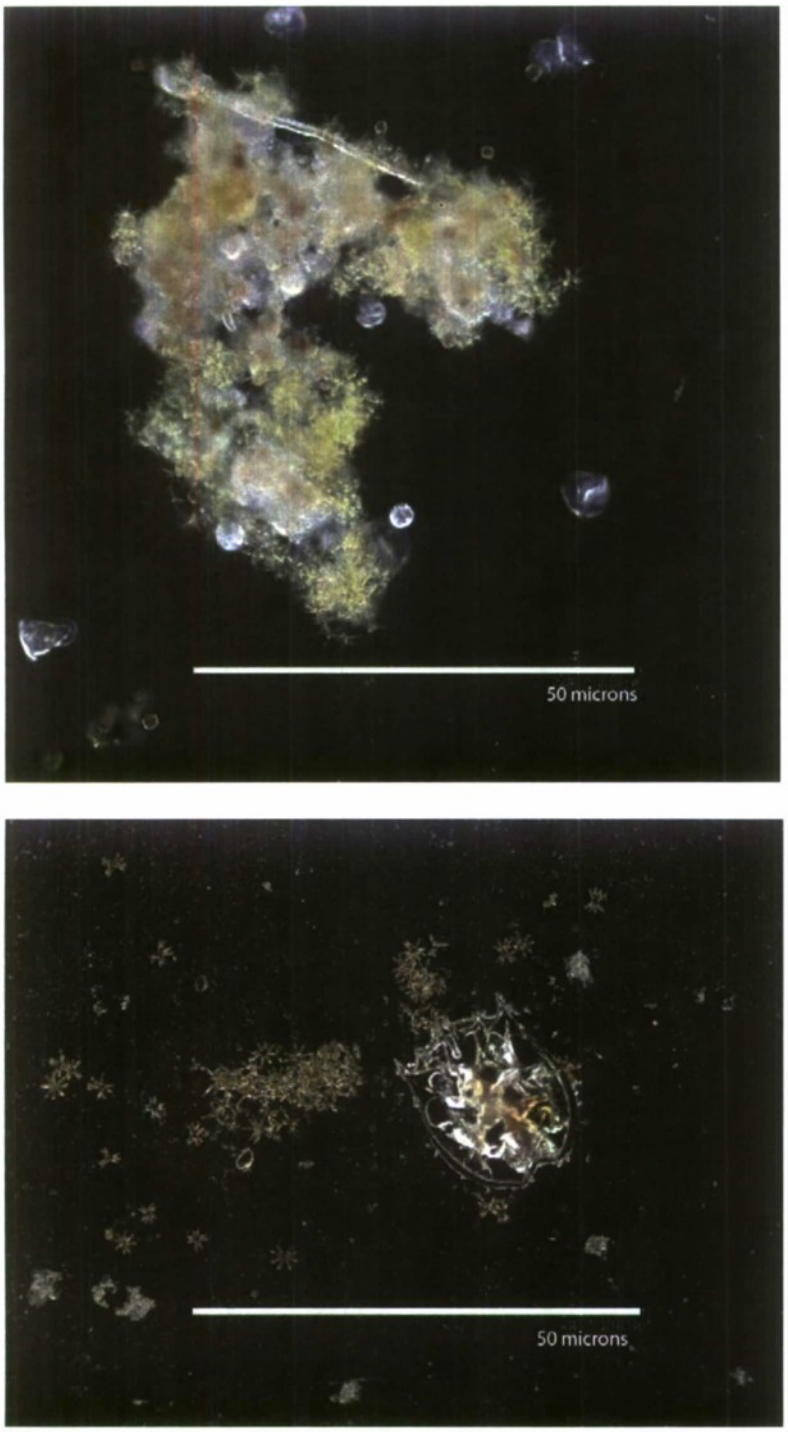


Figure 3-17. Micrographs of inoculated NW pond sample (top) and control NE pond sample (bottom). The inoculated ponds displayed algal-bacterial flocs and the presence of flocculation-promoting organisms such as stalked ciliates (light-colored shapes at a distance from the floc). The control ponds were often dominated by poorly flocculated cells. The lower image also shows a decaying rotifer body.

3.6 Extent of In-pond Settling

If a significant mass of algae settled within the ponds, then the biomass production of the ponds would be under-estimated based on the effluent solids concentrations. Accumulation of settled algae within the ponds was monitored in case the mixing speed was low enough to allow extensive settling. A sludge judge was used to measure sludge depth at various point in the ponds, as described in the Materials and Methods section. In high-rate ponds, settled algae accumulate in eddy zones such as the corners of the tanks or near the ends of the central baffles. Maximum accumulation in these eddy zones was 5-15 cm of solids depth, and that depth did not increase, indicating steady state accumulation. Other areas of the ponds had no measurable solids accumulation. Thus, the effluent volatile suspended solids mass flow was a good indicator of solids production.

3.7 Dissolved Oxygen

In the California climate, high-rate ponds treating sewage have design residence times of 3-10 days, depending on season and treatment objectives. The 3-d hydraulic residence time (HRT) used during 2007 was selected in order to test the most land-efficient HRT and to provide a high SBOD loading. A problem occurred with the 3-d HRT, however. In diel measurements, we found that dissolved oxygen (DO) levels decreased from 2-4 mg/L during the day to <0.1 mg/L at night. Although no malodor developed, these low DO concentrations are an odor risk. Remedial steps include lengthening the HRT, decreasing the pond depth, and nighttime aeration. For Experiment II, we chose to lengthen the HRT to 5 days by decreasing the inflow rate.

3.8 Experiment II Preliminary Results

The Experiment I data presented in this report, including the above, were collected during operation with a 3-d hydraulic residence time. The longer residence of 5 d is currently being tested at the pilot facility. The results so far are quite different than that of the 3-d residence time—in terms of higher DO concentrations (3-7 mg/L in the day) but, more dramatically, in terms of flocculation. The control ponds have formed larger flocs (3-15 mm), but they are not copious enough to adsorb much of the colloidal material. These flocs tend to be comprised of short filamentous algae strands, to be identified. The inoculated ponds are fully flocculated with large (3-5 mm) nonfilamentous flocs. These flocs settle almost completely within 10 minutes, and the supernatant is clear with <8 mg/L TSS remaining (**Figure 3-18**). This excellent flocculation has been consistently achieved in the inoculated ponds from late January to the present (late February). Influent water quality and weather conditions are not especially different than December, so it is likely that the increased average cell age of 5 days has promoted the better flocculation. Under the microscope, the flocs are seen to be heavily colonized by stalked ciliates and other organisms, as described above. Studies on the high-rate ponds and lab cultures are continuing with 2008 C3RP support. The new work will include continuous monitoring of DO and temperature using autosampler devices purchased with a $18,000 grant from the Civil and Environmental Engineering Department Student Fee Initiative Committee.

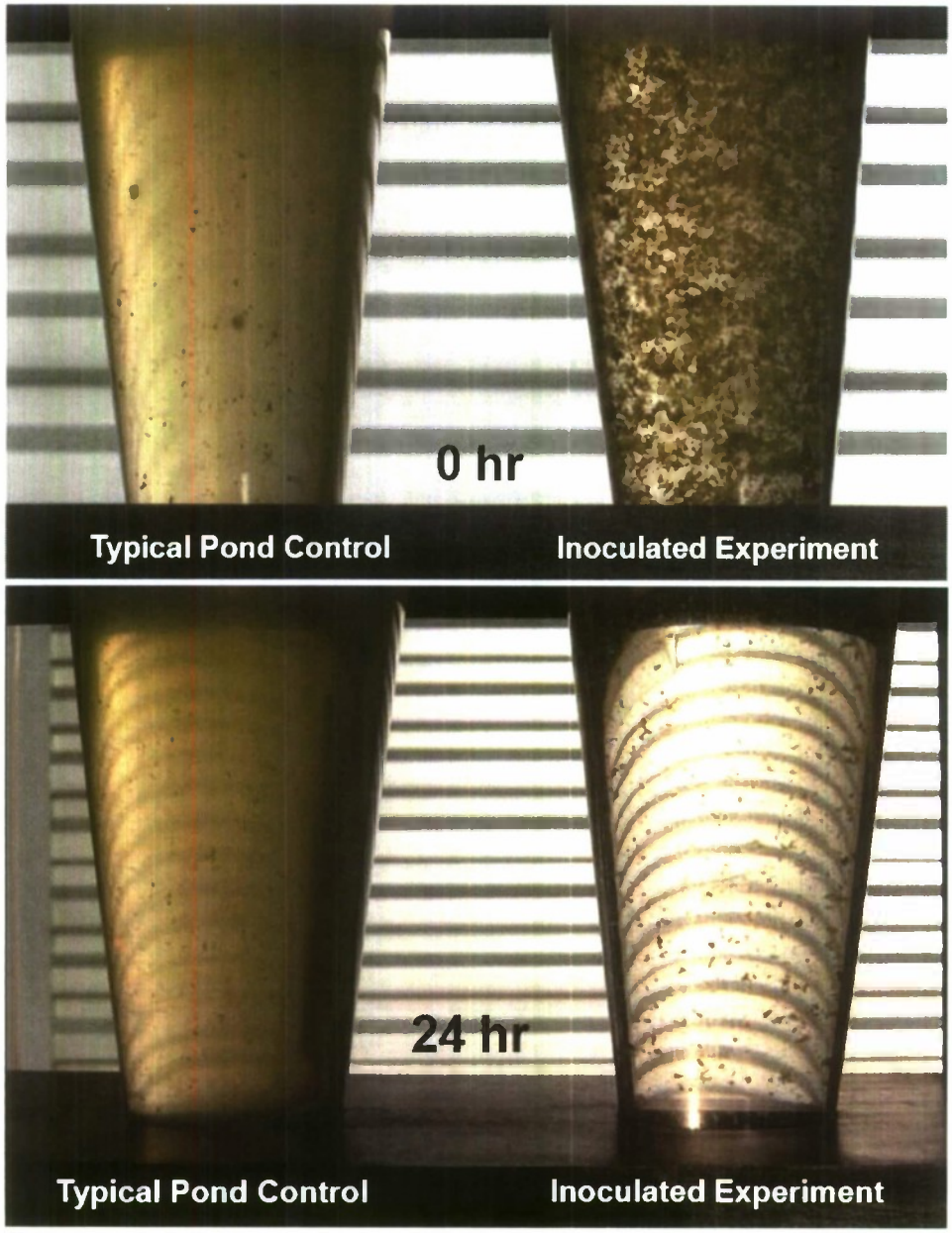


Figure 3-18. Bioflocculation and settling in Imhoff cones during Experiment II. TSS concentrations during Experiment II have improved to <8 mg/L for the inoculated ponds after 24 hours of settling. Nearly compete settling now occurs within 20 minutes. All the ponds are being operated with a 5-day hydraulic residence time during Experiment II, which is ongoing.

SECTION 4: CONCLUSIONS

During 2007 and early 2008, a considerable number of accomplishments have been made in our efforts to complete the development of simple, low-cost methods for wastewater reclamation and algae production. The lack of an affordable method to harvest microalgae has been a major impediment to both pond-based wastewater treatment and algae biofuel concepts for many years. Several hypotheses or goals related to algae sedimentation and treatment were presented in the original 2006 C3RP proposal.

Hypothesis 1: The total suspended solids concentration in settling tank effluents (TSS_S) is a function of the ratio of soluble biochemical oxygen demand concentration in the influent to the algae treatment pond ($SBOD_{in}$) and the total suspended solids concentration in the treatment pond (TSS_{pond}) but not a direct function of $SBOD_{in}$ or TSS_{pond} individually.

Hypothesis 2: TSS_S can be maintained consistently below the typically regulatory limit of 40 mg/L through manipulation of $SBOD_{in}/TSS_{pond}$.

Hypothesis 3: The settling tank effluent total BOD can be maintained consistently below the typically regulatory limit of 40 mg/L without increasing the treatment pond depth or hydraulic residence time outside the typical practical ranges of 0.3–1 m deep and 3–7 days, respectively.

To test these hypotheses, we constructed a pilot plant at a municipal wastewater treatment plant where sewage, partially-treated to various levels, was available for the pilot plant feed. The pilot plant included four tanks to represent high-rate ponds that could be individually operated. This replication allowed controlled experiments. Not enough data has been collected yet to statistically test these hypotheses, so preliminary results are described below.

Hypothesis 1 held but only up to a SBOD/TSS ratio of about 0.42. Ratios of 0.33-0.42 resulted in average settled TSS concentrations below the 40 mg/L target. At higher ratios of 0.6-1.5, the settled TSS concentrations were 40-60 mg/L. Additional work is needed to learn if these effects will hold during the various seasons.

The correlation of settled TSS to SBOD and pond TSS was present in the current studies, but that could have been coincidental to declining pond TSS concentrations as winter approached. Collection of more data during the various seasons should provide the information needed to test the independence of TSS and SBOD separately from the ratio of the two.

As described above, Hypothesis 2, appears to be true, but more data needs to be gathered to build confidence in our ability to maintain settled TSS below 40 mg/L. Continuous flow settling units are under construction to better simulate conditions at full-scale pond systems.

Hypothesis 3 was not found to be true. Although effluent SBOD concentrations were below the 40 mg/L target, with the addition of the particulate BOD caused by the residual 20-40 mg/L of algal VSS, the total BOD concentrations would have exceeded 40 mg/L. The low, <8 mg/L TSS concentrations being achieved during early 2008 indicate that the goal of Hypothesis 3 is achievable.

The mechanisms behind the improved settling seen in 2007-2008 have also been explored. Microscopic examination of the cultures suggests that higher organisms such as ciliates, common to activated sludge, are promoting the bioflocculation of the inoculated ponds. Although the control ponds have achieved lower settled TSS concentrations than full-scale high-rate pond systems, activated sludge organisms did not thrive in the control ponds. Instead, bacterially-mediated flocculation may have been the major mechanism of bioflocculation in the absence of routine inoculation with activated sludge organisms.

Beside the basic water quality parameters of TSS and BOD, we also monitored ammonium-nitrogen uptake by the pond cultures. Mean ammonia removals ranged from 40-64% without consistent differences among the control and inoculated ponds. Ammonium removal is a particularly energy-intensive process in conventional wastewater treatment technologies, so the removals attained by the ponds are beneficial. More complete ammonium removal could be obtained with longer hydraulic residence times (HRTs) or with addition of carbon dioxide.

During 2008, the main topics to be studied further are the reliability of bioflocculation, BOD removal at longer HRTs, maintenance of nighttime dissolved oxygen, ammonium removal, and resolution of the low productivity of the control ponds.

SECTION 5: REFERENCES

APHA (2005). *Standard Methods for the Examination of Water and Wastewater*, 21st ed., American Public Health Association, American Water Works Association, and Water Environment Federation.

ASCE (2005). "Report card for America's infrastructure," American Society of Civil Engineers, www.asce.org/reportcard/2005/page.cfm?id=35#one, accessed 11/30/05.

BEI, Inc. and EOA, Inc. (1988). City of Sunnyvale Algae Digestion Study, Bracewell Engineering, Inc. and EOA, Inc. August 1988, pp. 41.

Bryan, T. (2006). "A Call to Action," *Biodiesel Magazine*, February issue.

Crites, R., and G. Tchobanoglous (1998) *Small and Decentralized Wastewater Management Systems*, WCB/McGraw-Hill, New York.

DOD (2003). *Base Structure Report*, Fiscal Year 2003, Department of Defense, Office of the Undersecretary of Defense (Installations).

Downing, J.B., E. Bracco, F.B. Green, A.Y. Ku, T.J. Lundquist, I.X. Zubieta, and W.J. Oswald (2002). "Low cost wastewater reclamation using the Advanced Integrated Wastewater Pond System Technology and reverse osmosis," *Water Science and Technology*, Vol. 45, No. 1, pp. 117-125.

Eisenberg, D.M. (1981). *Wastewater Treatment, Energy production, and Energy Conservation in an Algal-Bacterial System*, Ph. D. dissertation, Department of Civil Engineering, University of California, Berkeley, pp. 289.

Feffer, A.S. (2007). *Effect of CO2 Addition on Algae Wastewater Treatment*, MS thesis, Civil and Environmental Engineering Department, California Polytechnic State University, pp. 160.

Fremming, Parson, and Pecchenino, Inc. (2005). Personal communication from L. Fremming, based on experience at the Delhi and Hilmar, Calif. AIWPS® facilities.

Frost, D. (2008). MS thesis in preparation, Civil and Environmental Engineering Department, California Polytechnic State University.

Garcia, J., F.B. Green, T.J. Lundquist, et al. (2006). "Long term diurnal variations in contaminant removal in high rate ponds treating urban wastewater," *Bioresource Technology*, Vol. 97, No. 14, pp. 1709-1715.

Green, F.B., T.J. Lundquist, and R.E. Brown (2006). "Energy Efficient Wastewater Treatment at Delhi and Hilmar, California: Case Studies," paper presented at the American Council for an Energy-Efficient Economy Summer Study, August 13–18, 2006, Pacific Grove, California, pp. 8.

Green, F.B., T.J. Lundquist and W.J. Oswald (1996). *The Capture and Utilization of Methane from Advanced Integrated Wastewater Pond Systems*, Final Report to the California Energy Commission, Sacramento, California, Environmental Engineering and Health Sciences Laboratory Report No. 96-1, pp. 266 plus appendices, pp. 198.

Green, F.B., T.J. Lundquist, and W.J. Oswald (1995). "Energetics of advanced integrated wastewater pond systems," *Water Science and Technology*, Vol. 31, No. 12, pp. 9-20.

Kirkpatrick, D. (2007). Announcement presentation for DARPA BAA07-08, Algae Biomass

Summit, November 16, 2007, San Francisco.

Lundquist, T.J. (2006). *A Reduction Pond for Denitrification of Agricultural Drainage*, Ph.D. dissertation, Department of Civil and Environmental Engineering, University of California, Berkeley.

Lundquist, T.J., F.B. Green and W.J. Oswald (2004). *Tertiary Upgrade Pilot Study Report*, Oswald Engineering Associates, Inc. prepared for the City of St. Helena, California, pp. 50.

Metcalf and Eddy (2002). *Wastewater Engineering: Treatment and Reuse*, 4th ed., McGraw-Hill Science/Engineering.

MWH (2005). Water Treatment: Principles and Practice, 2nd ed., Wiley and Company, Hoboken, New Jersey, pp. 1948.

Nurdogan Y. and Oswald, W.J. (1995). "Enhanced Nutrient Removal in High-rate Ponds," *Water Science and Technology*, Vol. 31, No. 12, pp. 33-43.

Oswald, W.J. (1990). "Advanced Integrated Wastewater Pond Systems," In: *Supplying Water and Saving the Environment for Six Billion People, Proceedings of the 1990 ASCE Env. Eng. Div. Conf.*, eds. U.P. Singh and O.J. Helweg, Amer. Soc. Civil Eng., 345 East 47th St., New York, NY 10017-2398, (November).

Patzek, T.W. (2005). "Thermodynamics of the corn-ethanol biofuel cycle," *Critical Reviews in Plant Sciences*, 23(6):519-567 (2004); web update 2005.

SBW Consulting, Inc. (2002). *Energy Benchmarking Secondary Wastewater Treatment and Ultraviolet Disinfection Processes at Various Municipal Wastewater Treatment Facilities*, San Francisco, CA: Pacific Gas and Electric Company, pp, 14, plus four attachments.

Scholze, R.J., and H.H. Zaghloul (2001). *Assessment of SCADA Technology Applications to Automate U.S. Army Water and Wastewater Sanitary Systems*, Report number: A037893, Engineering Research and Development Ceneter, Champaign, Illinois, pp. 28.

USACE (1999). *Design, Construction, and Operation - Small Wastewater Systems*, Publication No. EM 1110-2-501, U.S. Army Corps of Engineers.

USEPA (2007). *Waste Stabilization Pond Design Manual*, in preparation, Office of Water, Office of Research & Development, Cincinnati, Ohio.

WEF (1997). *Energy Conservation in Wastewater Treatment Facilities*, WEF Manual of Practice No. FD-2, Water Environment Federation, Alexandria, Virginia, pp. 190.

Weissman, J.C., R.P. Goebel, J.R. Benemann (1988). "Photobioreactor design: Mixing, carbon utilization, and oxygen accumulation," *Biotechnology and Bioengineering*, Vol. 31, No. 4, pp. 336-344.

WIN (2000). *Clean & Safe Water for the 21st Century*, Water Infrastructure Network, Washington, D.C., http://www.nacwa.org/getfile.cfm?FN=winreport2000.pdf.

Woertz, I., T. Lundquist, A. Feffer, and Y. Nelson (2008). "Lipid production by algae grown for treatment of dairy and municipal wastewater," submitted to the *Journal of Environmental Engineering*, American Society of Civil Engineers.

Circulation Control Wing Optimization Study for ESTOL Applications

Project Investigator:

David Marshall
Aerospace Engineering
California Polytechnic State University
San Luis Obispo, CA

Circulation Control Wing Optimization Study for ESTOL Applications, David D. Marshall

Objective

This study intends to extend the existing work into the use of circulation control techniques to find an optimal configuration for takeoff and landing of a nominal aireraft. This will require the development of an interface between the optimization techniques and the CFD tool, Fluent, in order to automate the search in the design spaee. Another development that will need to be aehieved is a database interface into the CFD results such that the design space data can be searched and queried during the analysis phase of the work. Other areas of possible investigation are the use of leading edge blowing to further the improve, sensitivity analysis on the impact of turbulenee models on the results, improved modeling of the engine power system and the aerodynamics and pulsed blowing techniques to reduee the required mass flow rates.

Task: CFD Analysis

One of the major drawbaeks for the eurrent simplified methods for ealculating takeoff distance is their lack of knowledge of how the lift and drag of an airplane vary throughout the takeoff run. Although the simplified models provide adequate results for traditional aircraft, the results for STOL aircraft are largely inadequate. This inadequacy is due to the fact that there are no analytieal methods to aeeurately calculate the lift and drag of an aircraft using powered lift such as circulation control. To better understand the aerodynamie forees throughout the takeoff process, CFD results from de la Montanya[3] were used to obtain the lift and drag for a given circulation control wing for a STOL passenger transport which can be seen in Fig(1) and Fig(2). The eirculation control wing which was analyzed uses a dual radius flap that extends 80% of the wing similar to that used by GTRI[4]. *Fluent*, a eommereially available CFD package, was used with the *Spalart-Allmaras* turbulenee model to obtain the lift and drag for various mass flows out of the ejection slot and at different angles of attaek. These preliminary results were then regression fit in order to implement them into the differential equations governing the takeoff.

Figure 1. STOL transport using Circulation Control[3]

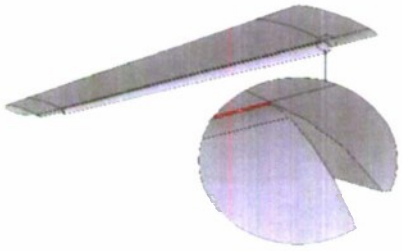


Figure 2. Circulation Control Wing[3]

Current CFD work has also included a study comparing different grids and how different turbulence models handle eireulation control. In order to accomplish this, multiple two-dimensional grids were created around the circulation control airfoil which were then solved under the same eonditions as those found in experimental data from Robert Englar which were reproduced in reference 5. the y+ values for the fine and eoarse grids were 50 and 500 respeetively and the *Spalart-Allmaras* and K-ω turbulenee models. The lift as a function of blowing coefficient for the diffcrent grids and turbulence models can be seen in Fig(3). The coarseness of the grid and the type of turbulence model clearly has a significant impaet on the results. The solution that most agrees with the experimental data utilized the coarse grid with the *Spalart-Allmaras* turbulence model. As a result of this study, the future 3-D solutions will be run using the *Spalart-Allmaras* model.

Full 3-D results are desired at a variety of diffcrent flap defleetions in order to develop a model that can aeeurately prediet takeoff performance. *Gambit*, which is the CFD gridding software included with

Fluent, does not easily allow for geometry changes like a flap defection. As a result, a semi-automated process was developed to create a grid around the given wing from Fig(2) with an arbitrary flap defleetion and example of which can be seen in Fig(4). Due to *Gambit's* inability to recognize similar edges when the flap is deflected at different angles, the process is required to be monitored and altered several times throughout the gridding process to achieve the desired results. Unlike in two dimensions, three dimensions poses a serious problem for computational time. The number of cells can quickly increase from 10-15 thousand in 2-D to upwards of 3 million in 3-D. In order to decrease computational time, several different solution techniques were investigated.

A number of different techniques were used decrease convergenee time for the 3-D grids. First, the grid was coarsened so that only the minimum amount of eells remained in order to provide a laminar solution. This decreased the cell count from over three million to somewhere around 750,000 depending on the flap defleetion. Although the boundary layer was not captured at all with the resulting y+ values upwards of 8000, the boundary layer could be resolved with the use of grid adaption. *Fluent* allows for many different types of grid adaption in order to resolve local flow effects in unstructured grids by dividing the cells in the desired location until the desired resolution is obtained. By adapting the grid with respect to y+ after it had converged on the coarse grid, *Fluent* was able to lower the y+ significantly and re-converge with a minimal number of additional iterations. An example of the adaption on the wing surface at the dual radius flap can

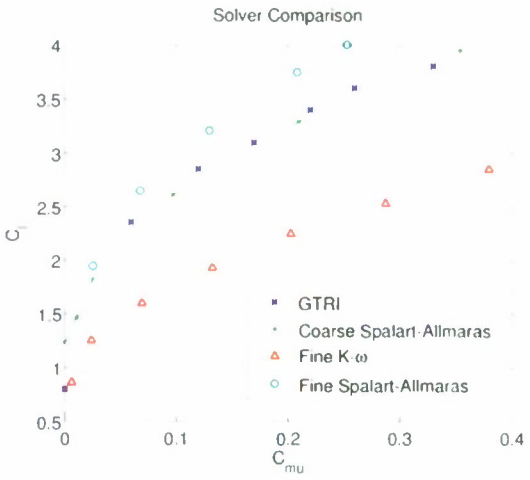


Figure 3. Solver Comparison

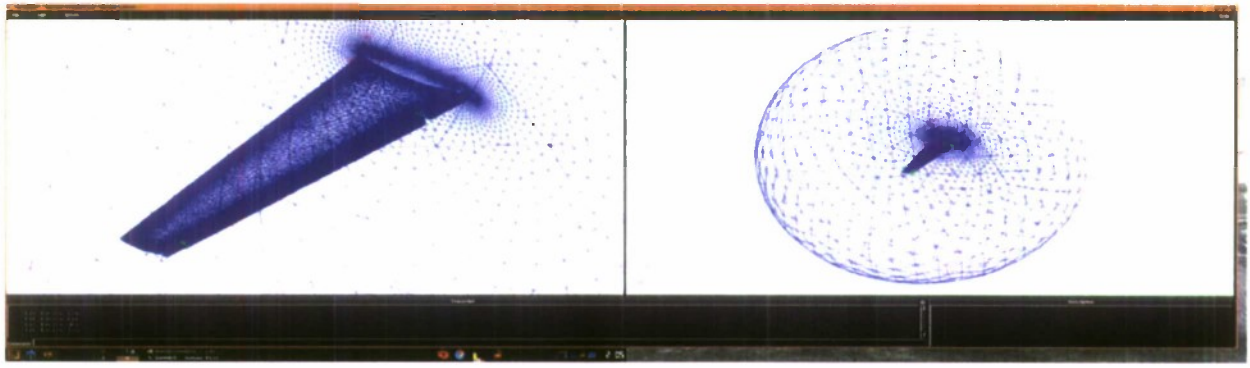


Figure 4. 3-D Grids Generated around the Circulation Control Wing

be seen in Fig(5). Preliminary results show that a series of 5 adaptions can reduce the y+ to a value of 570 which allows for the effects of the boundary layer to be included in the flowfield. Although the results of the adapted grid were looking promising, there was still an issue with computation time. The grid quickly inereased in size to over 3 million cells after a couple adaptions which drastically inereased convergenee time.

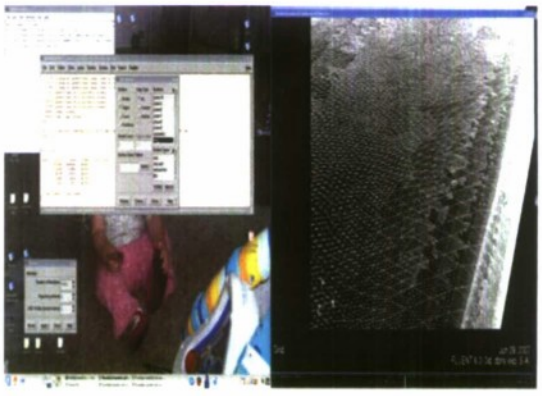


Figure 5. Y+ Adaption

The next technique to deerease divergenee time whieh was investigated was the use of multi-gridding. Multi-gridding is a proeess of using suceessively eoarser grids in order to eliminate low frequency error from the system. This ability is built into *Fluent* and, when enabled, allows the user to generate a speeified number of eoarser grids and eontrol over many other properties related to the elimination of error from the system. Although sweeping through a series of grids proved to take additional time for eaeh iteration, it was found that the number of iterations required to reach eonvergence was an order of magnitude lower when multi-gridding was used when eompared to a single grid. One of the drawbaeks known to exist with multi-gridding is the loss of non-linear flow effeets due to the interpolation between grids. To determine whether the results for eireulation control wing were being altered by the multi-gridding, a ease was solved with and without the use of multi-gridding and the resulting C_L was found to be only 1.6% different. Finally, the proper time to implement multi-gridding in order to minimize eonvergenee time was investigated. It was thought that a solution might be obtained quieker if a laminar ease was first eonverged and then the turbulence model was to be turned on. However, this was found not to be the ease. The minimum convergence time was found to oeeur with multi-gridding and the *Spalart-Allmaras* turbulenee model enabled throughout the case. It was found that before adaption, eonvergenee oeeurred in under 4 hours with 500 iterations. Once adaption was enabled, the time to iterate signifieantly began to increase which resulted in an additional 8.65 hours to iterate an addition 150 times to reaeh eonvergenee. A summary of the solver techniques and their performanee can be seen in Table(1) for a case with 90° flap deflection, Maeh number of 0.0836, angle of attaek of 0°, and a jet slot mass flow of 10 kg/s. The pathlines of the flow exiting the ejeetion slot for the *Spalart-Allmaras* ease with multi-gridding and 3 adaptions ean be seen in Fig(6).

Table 1. Solver Performance Comparison

Method	# Iterations to Convergence	Run Time (hrs)	Adaption(#)	Y+	CL	CD	CM
Spalart-Allmaras with multi-grid at 2500	3000	15.88	no	8202	2.6234	0.4732	2.0722
Spalart-Allmaras with full multi-grid	650	12.4	yes(3)	638	2.5830	0.4901	2.0450
Spalart-Allmaras with full multi-grid	500	3.75	no	8202	2.5980	0.4700	2.0534
Laminar to Spalart-Allmaras	2000	4	no	9000	2.5550	0.4472	2.0187

Future CFD work will include using the semi-automated gridding proeess to ereate additional grids in order to adequately resolve the design spaee of 0-90° flap defleetion. Additional time will be spent attempting to further lower the y+ value to a target of 250 to ensure the effeets of the boundary layer are fully influeneing the flowfield. Additional work will also inelude adding ground effect to the CFD models to more aeeurately prediet the lift and drag throughout the ground run. Finally, the

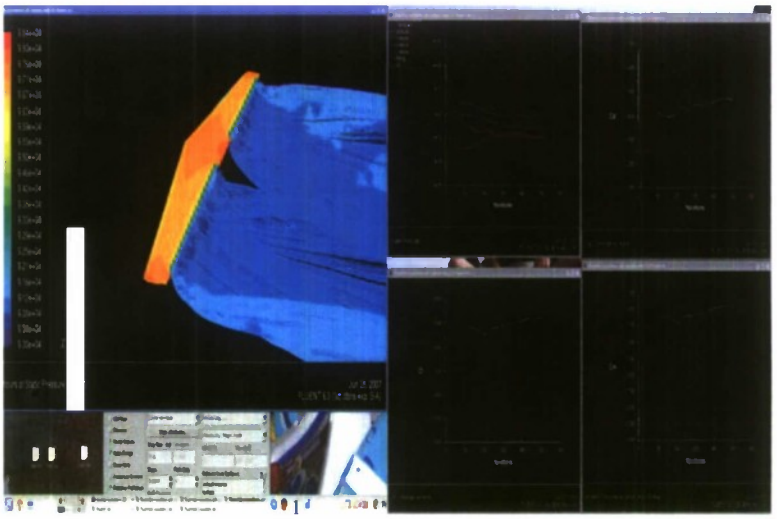


Figure 6. Pathlines from the Jet Slot

effect whieh the one engine inoperative eondition has on a cireulation control wing will be analyzed in order to preeisely prediet the balaneed field length for an aireraft using cireulation control. This situation will be of primary interest due to the strong eoupling between the powered lift and the thrust of the aircraft. All of these CFD results are stored within a single database sueh that any combination of eolleeted data ean be retrieved and a meta model of the data ean be ereated.

Task: Database Layout

The database used for this researeh was a MySQL database with access provided through a text based user interfaee written in shell seript and Perl. Once an experimenter has completed a eomputer simulation, they run the seript, whieh finds and stores meta model parameters in the database. Currently the interfaee pulls out all the data needed from the Fluent files and future work will store data required to rerun the case. Fluent input parameters and solution data (sueh as CL, CD, CM and slot jet velocity) are all stored and updated in the database. Eaeh experimental ease input is traeked by a unique experiment id, whieh links the data to experiment parameters. The nature of the MySQL database prevents duplicate storage of the same data to insure that only the most relevant data is available for export. To ereate a meta model, the data from the database is exported using seripts to a Matlab file that ean then be input into the response surfaee modeling tool. To use a different model only the output script need be modified, future work will inelude a number of different modeling options in the interfaee along with the option for ereating a eustom output format.

An advantage of the MySQL database and seript eombination is that it is robust and aets as a data hub for any remote Linux or Unix maehine. This provides experimenters in different loeations the ability to upload data to the database and have it meta modeled remotely. The results of this meta modeling ean then be uploaded to the database and any user ean aeeess that data. This is particularly benefieial in the ease of CFD beeause large eomputational eases ean be run on eomputers at different locations and the data ean be integrated seamlessly. If the Linux system were overly restrictive to users, it is possible to port the seripts over to other operating systems or to ereate a web based GUI with additional work.

Task: Meta Modeling

While modeling teehniques attempt to prediet the behavior of a system, meta models attempt to prediet the behavior of a model. This is useful for models that are expensive to obtain data points, sueh as CFD experiments, beeause relatively few points ean be used to build a meta model of the entire design region. It is impraetieal to ealeulate design points for the entire design spaee for CFD situations beeause takeoff ealeulations are based on a number of ehanging parameters, and ehanging any one parameter requires a new solution that will take at least three hours for the ease outlined above.

Using Matlab's response surfaee tool, a full quadratie model was fit to the input data points. Lift and drag eoeffieient data was taken from the Fluent runs of the aireraft and related to the blowing eoeffieient. The blowing eoeffieient was ehosen as the eommon input variable for two reasons. First, it is a funetion of the ineoming veloeity, whieh is the integrating term for the ground roll equation. Seeondly, the blowing eoeffieient eontains the mass flow rate term, whieh represents the exhaust gas velocity and ean be used to model engine out situations. For simplieity in this early study, the mass flow rate was set as a eonstant for the length of a ground run. This was a reasonable beeause the exit velocity and density of the flow exiting the baek of the eireulation control flap ehanges only slightly over the eourse of the takeoff as long as all engines are functioning. Future versions of the model will include the eoupling for OEI and for the nominal eonfiguration.

The full quadratie model used to represent the lift and drag eoeffieients ineluded the eoupling

terms between angle of attack and blowing coefficient. The equations for lift coefficient and blowing coefficient are shown below:

$$C_L = 1.08 + 11.7C_\mu + 0.0809\ \alpha - 0.0116\ C_\mu\alpha - 18.07\ C_\mu^2 - 0.0031\ \alpha^2 \quad (1)$$

$$C_D = 0.0982 + 2.79\,C_\mu + 0.0133\ \alpha + 0.0455\ C_\mu\alpha - 4.20\,C_\mu^2 - 0.0002\ \alpha^2 \quad (2)$$

The root mean square errors for these two functions were 0.242 and 0.0565 for lift and drag respectively. While this represents a large percentage currently, additional data points will be chosen that reduce this error and in the final automated process additional data will be requested if the error exceeds a specified threshold.

In addition to response surface models, other models are currently being investigated that can be used to represent the takeoff requirements of a circulation control wing. Due to the non-linear nature of the coupling between systems, non-linear models present a natural choice for the application. Once an appropriate non-linear model has been chosen that fits the behavior of the system it will be implemented and compared to the response surface in order to obtain the best possible curve fit. Gaussian processes are also being researched in order to provide another option for meta modeling.

Gaussian processes provide the most robust way to deal with the experimental data points at the added cost of model-design complexity. Specifically, since Gaussian processes are dependent on a model function that is specified when designing the model, they can take forms that have the same trend as the model. For example, if the modeled function is a step function the Gaussian process can be made to favor that type of behavior. Gaussian processes can also behave like a regression or interpolation, which can be optimized for the problem at hand. The optimization insures helps to create the best model. Ideally, all three methods will be implemented and results compared to determine the best choice for speed, accuracy, and ease of use.

Design of experiment (DOE) methodology is being investigated to insure that the least number of runs are used to obtain the desired errors. However, classic DOE methods do not apply to computer experiments because for a given set of inputs the outputs will always be the same, effectively eliminating the random error. The main source of error for such experiments comes from the bias error of the system and methods of reducing bias error are a main avenue of investigation[8].

Task: Performance Model

With the aerodynamic data obtained and lift and drag functions provided, the final step was to solve the equations of motion governing takeoff. The equation which governed the ground run was taken from Raymer[7] and is listed as Eq(3).

$$S_g = \int_0^{V_{LO}} \frac{d(V_\infty^2)}{\left(\frac{1}{m}\left(T - D - \mu_r(W - L)\right)\right)} \quad (3)$$

This integral was solved numerically by changing it to a differential equation and using an ordinary differential equation solver from MATLAB called ODE45. At this point of the analysis, the only parameter which was able to be varied with velocity was the blowing coefficient. From that blowing coefficient variation, the lift and

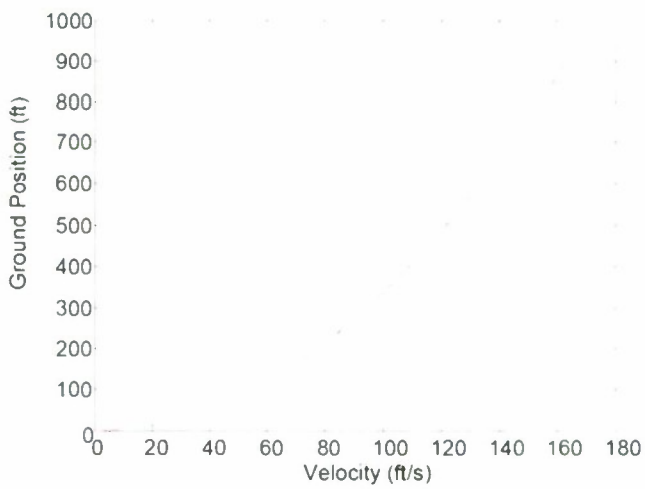


Figure 7. Ground Position as a Function of Velocity

drag were able to be determined throughout the takeoff process as described earlier. Within the ODE solver, those functions were entered in order to increase the accuracy of the model. Shown below in Fig(7) is the ground location versus the velocity of the aircraft and Fig(8) shows how the lift and drag varied with different mass flow rates (in slug/s) from the ejection slot. From Fig(8) it is clear that with a constant mass flow rate, the lift decreases significantly throughout the ground run due to the decreasing blowing coefficient. This is the exact opposite trend that is desired during a takeoff run since a high C_L at the start will increases the induced drag when high acceleration is desired. The optimal takeoff run would therefore have very little mass flow at the start to minimize drag and a lot of mass flow at the end to maximize lift. This relationship between the mass flow, blowing coefficient, and coupling that with the actual output of an engine will be examined in the future to come to a more complete understanding about how to provide the shortest takeoff distance possible.

With the current performance model, a preliminary analysis of the required takeoff field length was calculated. This includes the distance on the ground, transition, and the distance in the air required to clear a 50 ft obstacle. The total distance using this model was found to be 2,124 feet. This was then compared to the value obtained from Anderson's[8] preliminary design takeoff equation which yielded a value of 1,201 feet. The difference between the two lengths comes from the many assumptions that go into the takeoff equation, which include a constant C_L and C_D. This difference is very large and would cause a significant redesign if the aircraft was optimized towards the takeoff constraint early in the design. In the future, the takeoff performance model will couple in how thrust, C_L and C_D vary with Mach number, Reynolds number, blowing coefficient, angle of attack, and flap deflection. This will allow for a more accurate representation of how an aircraft with circulation control performs during takeoff.

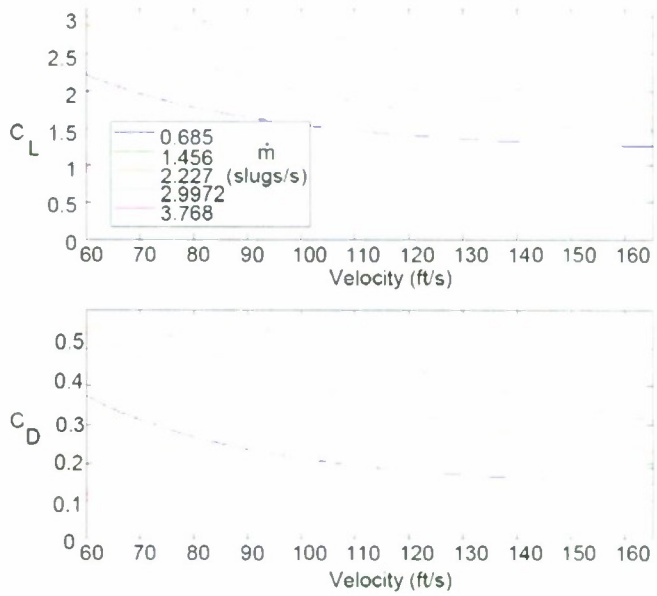


Figure 8. Lift and Drag vs. Velocity for variable mass flow rates

References

[1]Torenbeek, Egbert. *Synthesis of Subsonic Airplane Design.* Delft University Press, Delft, The Netherlands. Ch5 and App. K 1982.

[2]Bobbitt, Percey J., Margason, Richard. *Analysis of the Take-off and Landing of Powered Lift Aircraft.* American Institute of Aeronauties and Astronauties, Ine., 2007

[3]de la Montanya, Julianna Belle. *Circulation Control and Its Application to ESTOL Vehicles*, Master's Thesis, California Polytechnic State University, San Luis Obispo, CA. June 2006.

[4]Englar, Robert J. (Georgia Tech Researeh Inst., Atlanta);Smith, Marilyn J. (Georgia Teeh Researeh Inst., Atlanta);Kelley, Sean M. (Georgia Teeh Researeh Inst., Atlanta);Rover, Riehard C., III (Georgia Teeh Researeh Inst., Atlanta). *Application of circulation control to advanced subsonic transport aircraft. Part I – Airfoil Development.* Journal of Aireraft 1994 0021-8669 vol.31 no.5 (1169-1177)

[5]Liu, Yi, Sankar, Lakshmi N., Englar, Robert J., Ahuja, Krishan K. *Numerical Solutions of the Steady and Unsteady Aerodynamic Characterishtics of a Circulation Control Wing.* GTRI Report A5928/2003-1 App. B. Georgia Teeh Researeh Institute, GA 2003.

[6]Box, George E. P. *Reponse Surfaces, Mixtures, and Ridge Analyses.* Wiley-Interseience, New Jersey. Ch. 14, 15. 2007.

[7]Raymer, Daniel P. *Aircraft Design: A Conceptual Approach Fourth Edition.* Ameriean Institute of Aeronauties and Astronautics, Ine., Virginia, 2006

[8]Anderson, John, D. *Aircraft Peformance and Design.* MeGraw-Hill Seienee Engineering, New York. Ch 6. 1998.

Proposed Future Work Associated with the Circulation Control Wing Modeling Effort, David D. Marshall

The current work related to circulation control wings is developing an infrastructure capable of using CFD to model the complicated flow physics associated with this technology. Significant advances have been achieved with the CFD modeling such that trends are becoming apparent that will be useful in the design of aircraft that will utilize this technology. In addition, a database storage facility has been developed that allows the efficient storage and retrieval of a large number of datasets. It is planned to have the database capable of holding multiple records of the same configuration in order to be able to store multiple runs of the same experiment, or results with varying levels of fidelity. Finally we are currently developing a takeoff and landing simulation that is capable of utilizing the database, the CFD results and preliminary models of the propulsion and control systems. This tool will be useful to allow the aircraft designer to determine the takeoff and landing characteristics of their aircraft with more accuracy than the current popular takeoff and landing models. Thus at the end of the current work a significant portion of the architecture will be in place to facilitate the use of advanced aerodynamic concepts, such as circulation control wings, in the preliminary design of next generation aircraft. What follows is a list of possible future works associated with this work.

Experimental Validation – There exists little reliable, experimental validation of a wide variety of circulation control wing designs, and there is a need to perform some experimental validations. We have an established relationship with some of the wind tunnel operators at NASA Ames, and they have generously donated their facilities to run experiments that we developed for little cost. What we will need funding is the design and fabrication of models capable of generating validation data for the CFD results already obtained. In particular, there is interest in leading edge suction for boundary layer control and externally deployable ducts that act as inlets to the compressors that are accelerating the jet of air leaving the trailing edge. These concepts are exciting extensions to the basic circulation control wing concept, and are challenging concepts to model with the current state of the art in CFD. Developing reliable experimental data for these geometries will provide preliminary results as to the advantages of these new concepts. In addition this data will act as validation data for CFD analysis of these concepts.

Lower Fidelity Approximations – While CFD and experimental data can be obtained for circulation control wings, there is a significant amount of effort required to develop even one data point. This makes the use of CFD (and especially experimental techniques) impractical for doing design optimization or sensitivity analysis where thousands of cases need to be quickly evaluated. This highlights the need for lower fidelity models that have a quicker compute time at the expense of possibly lower accuracy. The proposed low fidelity model is to adapt standard panel methods that have a long history of producing accurate aerodynamic predictions with short computation times. A standard panel method solution technique would need to be modified to couple the boundary layer since it is the boundary layer that provides an integral role in maintaining the attached flow on the trailing edge Coanda surface. The boundary layer coupling via Integral Boundary Layer Methods or Parabolized Navier-Stokes equations can predict separation and provide reasonable estimates to the viscous drag. In addition the jet exit on the trailing edge flap needs to be modeled in order to account for the blowing coefficient variation in lift. It is not guaranteed that this approach will provide adequate results, but preliminary work has shown some promise in producing the qualitative aspects of the flow field by artificially moving the separation point of the flow in the trailing edge Coanda surface.

Database Generalization – The database that has been generated is a powerful extension to the typical methods that large collections of data are stored. Typically the data owner creates a complicated

directory strueture as a way of organizing the data. This has the obvious disadvantage of not allowing the results to be quiekly mined for a partieular subset of data. One example is the extraet all data with a lift eoeffieient between 0.9 and 1.2, an angle of attaeh between 1 and 3 degrees and the Spallart-Allmaras turbulenee model run. With the storage of the results into a database, this kind of information can be quiekly queried using SQL syntax. What would make the database even more useful is if the database were generalized sueh that experimental data, lower fidelity model results, and other general datasets could be collected.

The Cal Poly Spacecraft Attitude Dynamics Simulator

Project Investigator:

Eric Anthony Mehiel
Aerospace Engineering
California Polytechnic State University
San Luis Obispo, CA

Abstract

The Cal Poly Spaee Teehnologies and Applied Researeh (CP/STAR) Laboratory is currently engaged in the design and development of a ground based Spaeeeraft Attitude Dynamics Simulator, or CP/SADS. The CP/SADS is a four axis pyramidal formation set of reaetion wheels which rests on a spherieal air bearing. When the eenter of gravity of the reaetion wheel structure is eolloeated with the pivot point of the air bearing, no gravity torque is ineumbent on the CP/SADS. The resulting rotational dynamies are equivalent to the rotational motion of a satellite in earth orbit. The CP/SADS is reeonfigurable via software and hardware and is eapable of simulating various mass distributions. The CP/SADS acts as a test bed for spaeeeraft attitude control laws and spaeeeraft attitude sensors and aetuators. Under support of the C3RP program funded by the Office of Naval Researeh (ONR), the CP/SADS team has been performing basie development of the simulator. More speeifieally, the following projeets have been eompeted or are near completion. The Wireless Data Aequisition System (WDAS) required to free the CP/SADS from all eable torques has been designed and built in a breadboard fashion. Proof of eoneept eode for the WDAS operation has been written and aetual data and eommands have been aequired and generated respeetively. The CP/SADS Fine Balaneing System (FBS) have been designed and assembled and integrated with the strueture. Work to implement the FBS will eontinue the summer. Finally, a low eost star traeker based on eommereial off-the-shelf components has been integrated and eharaeterized.

Introduction

The Cal Poly Aerospaee Engineering Department is eurrently broadening the eurrieulum and research endeavors into the area of Space Systems Engineering. The eoneepts of design iteration and requirements development and flow are beeoming inereasingly important in today's aerospace industry. To further enhanee the eduеational goals of the department and to be involved in an area of signifieant applied researeh, our group has embarked on a quest to develop a spaee system simulation laboratory. Our initial development effort has been funded be the C3RP program.

During the past two years, with support from the C3RP program and the Office of Naval Researeh, our group has designed and built a four-axis, pyramid eonfiguration, reaetion wheel system to simulate spaeeeraft attitude control. Figure 1 shows a solid model of the system and a recent photo of the aetual hardware. The CP/SADS rests on a spherieal air bearing which provides 360 degrees of rotational freedom in the z-axis and +/- 35 degrees in the x- and y-axes. When the eenter of mass of the CP/SADS and the pivot point of the air bearing eoineide, no gravitational torque is felt by the platform and the miero gravity environment of space is effeetively reereated on the ground. When no gravitational torque in ineumbent on the CP/SADS, control laws developed for spaeeeraft attitude eontrol ean be implemented and studied. Further, attitude determination hardware and methodologies ean also be built and studied.

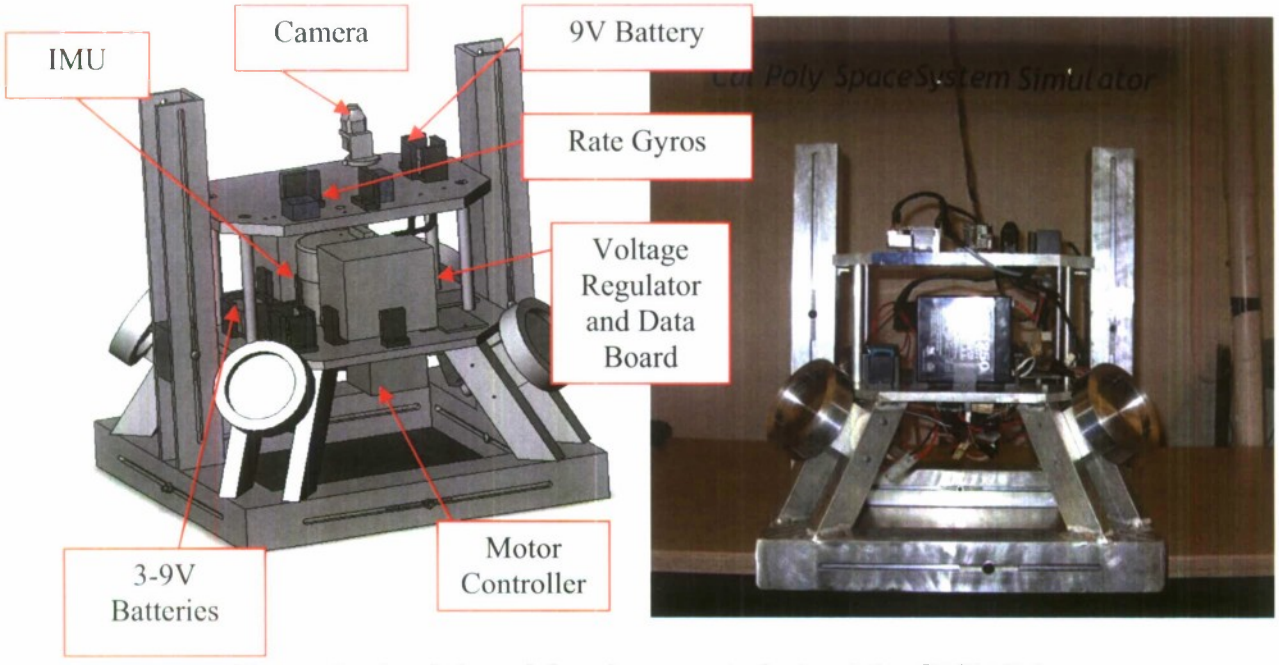


Figure 1. A solid model and a recent photo of the CP/SADS

Currently the CP/SADS has been tested about one axis and the results are promising. More refinement and development of the system will take plaee during the next several months prior to the end of the eurrent grant eyele. Specifically, our group has tested the system and proven that our design works and we have met all of our self-imposed performanee eriteria. For example, the CP/SADS is eapable of rotating at rates whieh are aetually faster than those originally proposed. To complete the proposed development of the CP/SADS, integration of various sensors and a completely wireless data aequisition system is neeessary.

Final Progress

Several projeets related to the CP/SADS were proposed during this round of C3RP funding. To demonstrate the preeision pointing ability of the DMRAC controlled CP/SADS, the final integration of the Wireless Data Aequisition System (WDAS) needs to be performed. A multidiseiplinary team of students has already been assembled for this task. The team, made up of Computer and Aerospaee Engineering students has performed the following tasks and written a Senior Projeet based on their work.

1. Integrate the Gumstix and Robostix into the existing CP/SADS test bed.
2. Design and write all required software for Gumstix and Robostix integration.
3. Develop and write a users manual describing the use of the Gumstix and Robostix.
4. Generate a senior projeet report detailing their results.

The WDAS has not yet been fully integrated onto the CD/SADS. Working with a student over the summer, we hope to eomplete the WDAS system and demonstrate the eommand and eontrol of the CP/SADS system about all three axes.

To date, signifieant progress has been made on the projeet items listed above. A group of seniors in the Computer Engineering Department designed and wrote eommunieation software eapable of:

- Establishing a network link with the Gumstix proeessor via the BlueTooth system,

- Reading data from the Rate Gyroscopes and Reaction Wheel Tachometers sensors, and
- Commanding the Reaction Wheel Speed and Direction.

Currently this software has been integrated into Matlab and Simulink for real time control by an Aerospace Engineering Graduate student. This student's thesis project involves the final implementation of the FBS. While not his work is not currently supported by this ONR grant, his work continues as he is nearing completion of the project and graduation. Also, an Aerospace Engineering MS student has been working to improve the sensing capability of the wheel speed sensors and provide angular acceleration data to the user. The final report from the CPE students and the Masters Thesis currently underway will serve as the user's guide to the Gumstix/Robostix WDAS with Matlab and Simulink.

Also, with the cooperation of Northrop Grumman through the Cal Poly Project Based Learning Institute, several students spent last year working on integrating the donated IMU onto the CP/SADS platform. Work in this area is ongoing. The challenge continues to be the wireless data acquisition requirement. Since most of these devices are not designed with wireless data systems in mind, the systems integration problem is significant.

Finally, our team was successful in implementing two iterations of a Star Tracking system for inertial attitude estimation of the CP/SADS platform. The first iteration of the Star Tracking system achieved reasonable results and is summarized in a Master's Thesis. However, after further review, a group of seniors took it upon themselves to implement a more sophisticated solution based on the QUEST algorithm. The results of this new design and analysis are shown in Figures 2 and 3.

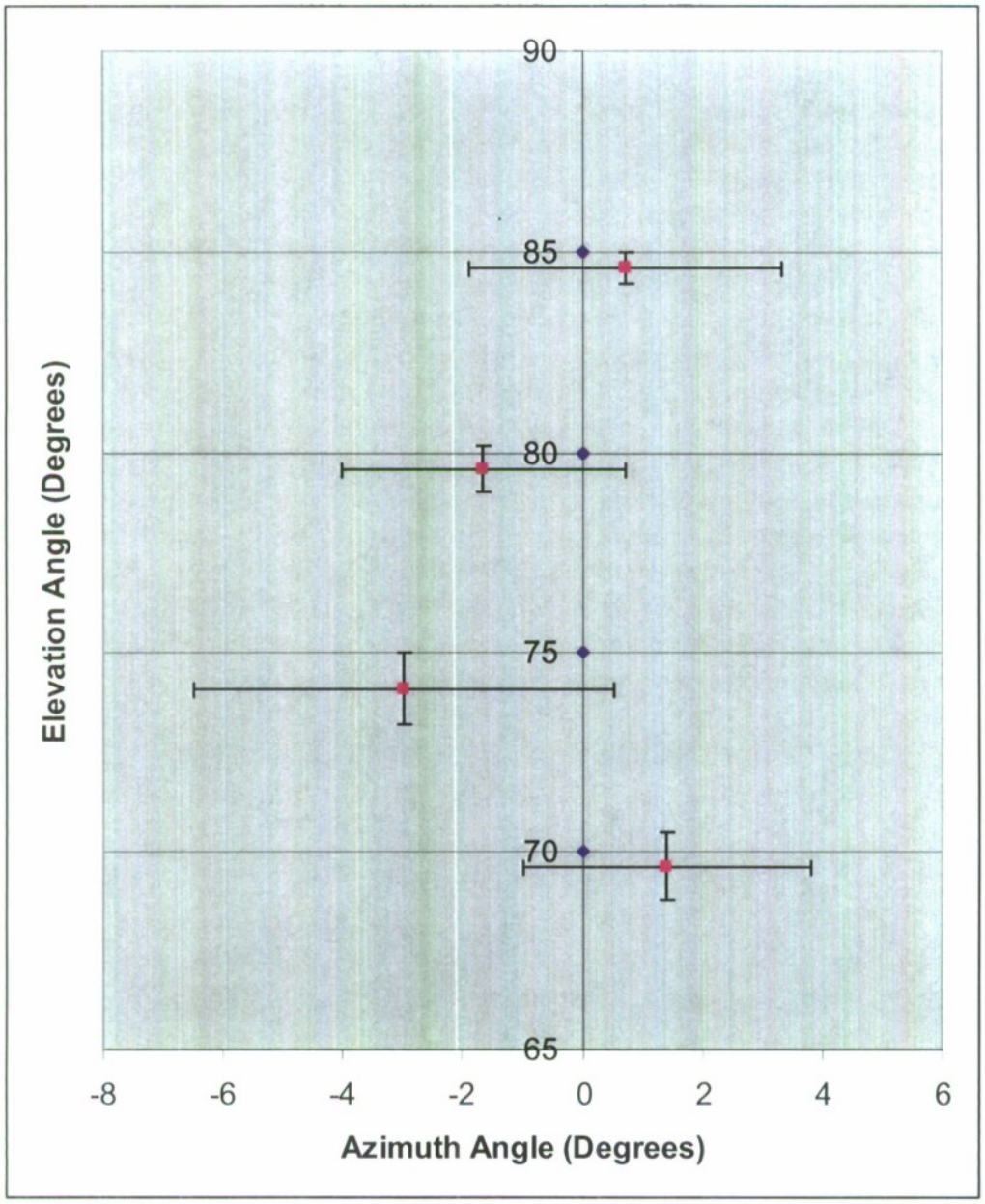


Figure 2. Experimental Results from QUEST based Star Tracker with variation in the Elevation Angle

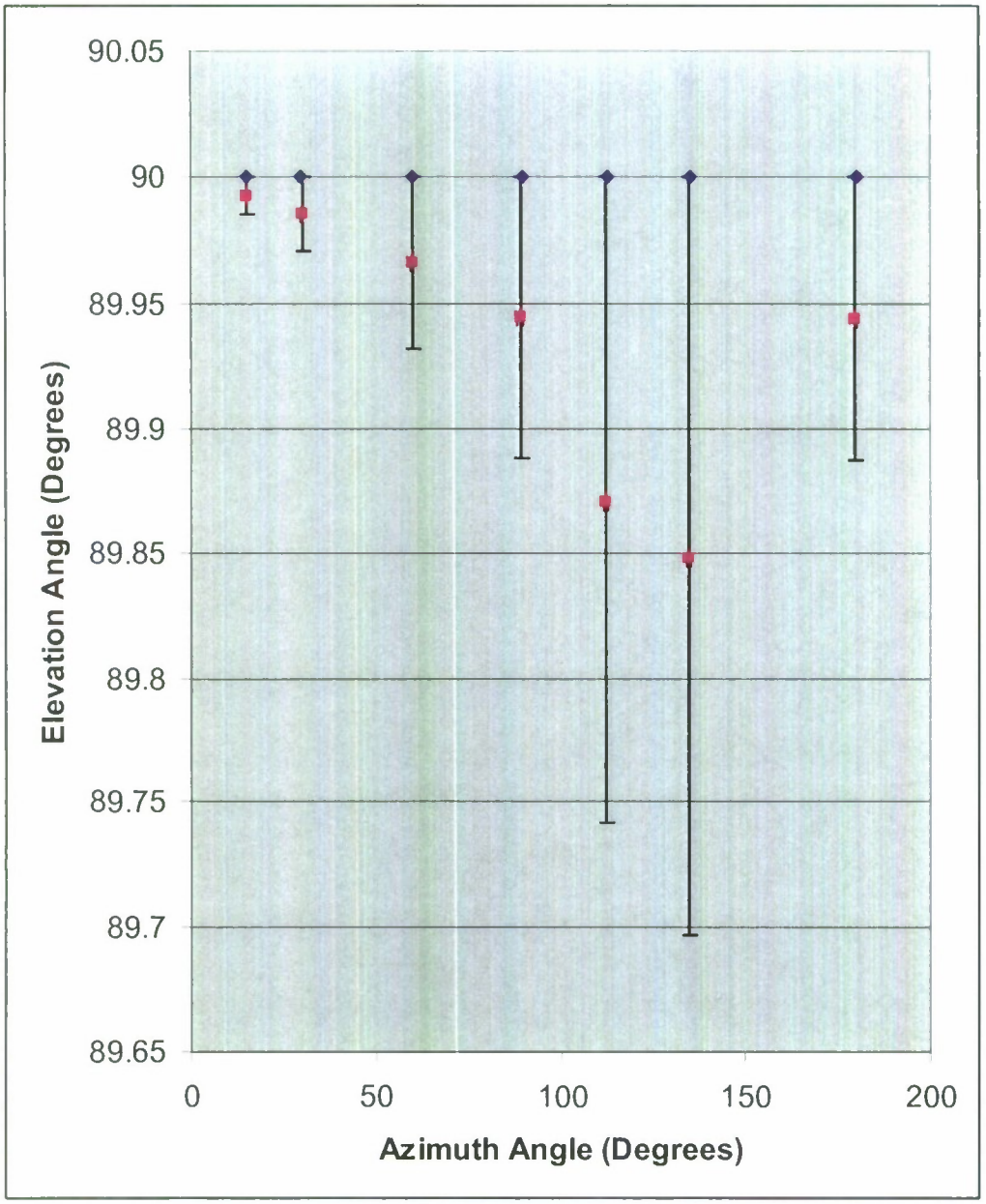


Figure 3. Experimental Results from QUEST based Star Tracker with variation in the Azimuth Angle

It is not total elear from Figures 2 and 3 because of the scales required to show the data, but the current Star Traeker design performs equally well in eaeh axis (Azimuth and Elevation); to within 1-2%.

In general we did not aehieve all of our goals as listed in the initial proposal. Speeifically, we have not finalized the System lD algorithm whieh is eritieal to the balaneing of the CP/SADS platform. However, work in this area is ongoing and we feel this project is nearing eompletion of a first iteration. Also, we did build two versions of a star tracking system for inertial attitude estimation, but we have not designed or built the required metrology suite to determine the validity of our attitude estimation results with respeet to inertial spaee.

Conclusions and Future Work

Overall we feel the C3RP project was a success. Significant progress was made to the overall goal of developing a working spacecraft simulator and control law test bed. To be sure we did not achieve all of our goals and more work needs to be accomplished. However, we have been successful in attracting sponsorship from other funding sources (Lockheed Martin, Northrop Grumman, and ATK) and we will continue to pursue our ultimate goal: to have a working spacecraft dynamics simulator.

In the future we plan to study the implementation of novel attitude control algorithms given the processing constraints of most small satellites. Specifically, can robust, adaptive control laws be used on board a small satellite without resorting to high bandwidth processors?

We are also interested in studying the ability of the CP/SADS to control flexible appendages mounted to the platform. The flexible appendages would simulate the structural disturbances caused by solar panels or other booms that operate on spacecraft. To achieve high accuracy pointing control the vibrational disturbances caused by the flexible appendages need to be rejected or accommodated. Certain adaptive control techniques look promising but have not been applied to these types of systems.

Finally, the CP/SADS provided both simulation and test capabilities of attitude control systems that is unique for a university environment. Our group would like to capitalize on our unique capability to procure other related projects and more funding. Currently we have submitted a proposal to Lockheed Martin to further support our project.

Development of Composite Material Solution for Advanced Nozzle Research

Project Investigators:

Joseph D. Mello, William Browning and Eric Peterson
Mechanical Engineering
California Polytechnic State University
San Luis Obispo, CA

Introduction and Objectives

The primary objective of this project was to use computational fluid dynamics and coupled thermal-structural finite element modeling to develop configurations and advanced material solutions for a concept annular aerospike hybrid rocket motor under development at Cal Poly by Carpenter, Murray, and Gerhardt. A secondary goal of the effort was to develop an experiment to complement the computational work and current rocket motor testing.

The work was carried out by two student investigators a graduate and an undergraduate, senior project student under the direction of the principal investigator. The students received all the funding in the effort. Their work resulted in a completed Master's thesis and a senior project which is near completion.

Preliminary Design Study of an Un-Cooled Annular Aerospike Nozzle

The major effort of this project was performed by William Browning a graduate student in mechanical engineering. This work is thoroughly documented in attachment 1 which is his master's thesis. A brief outline of the accomplishments and major findings follow.

Nozzle Geometry and Material design

A high-temperature material survey was performed and various aerospike configurations conceived with a variety of material solutions. Material vendors were located and consulted with regarding advanced refractory composites such as Silicon/Carbide and Carbon/Carbon. High-temperature metals were also examined.

Combustion and Heat Transfer Computational Fluid Dynamics, CFD Modeling

The methodology to develop estimates for gas side chamber and nozzle gas side heat transfer conditions were conceived and analytically tested. A combination of experimental data, semi-empirical analytical results and computational fluid dynamics were used to accomplish this task. This work culminated in surface film conditions as a function of axial location or station that could be used to asses aerospike designs and associated materials.

Heat Conduction Finite Element Model Development

Specialized FORTRAN user subroutines were developed and used to map the CFD film heat transfer conditions to an ABAQUS heat conduction finite element model. Transient and steady conditions were derived and used to examine temperatures for the array of aerospike material and geometric conditions conceived.

Thermal/Structural, Stress Model

The techniques to map temperatures or to couple the thermal results to an ABAQUS finite element stress model were developed as well. Stresses and

strains were then analyzed for the materials and geometries conceived. Special consideration was given to Carbon-Silicon Carbide Material and refractory metals.

Development of an Experimental Device to Examine Gas Side Heat Transfer Conditions

It was determined that the heat transfer conditions that drive this type of design should be confirmed experimentally. A senior project student Eric Peterson focused on this goal and has designed and fabricated a heat transfer measurement device. The solution for this device is a special rocket chamber-nozzle assembly which is to be used in the current hybrid motor test stand. This instrument is based on high-temperature metallic material and a series of thermocouples. Simple heat transfer and thermodynamic principles are to be used to estimate heat flux at the nozzle wall based on data collected from hybrid motor hot-fire tests.

Attachment 2 is a draft of the senior project report which details the design of this component.

Conclusions and Findings

Aerospike nozzles have been conceived and examined. This includes geometry considerations and a variety of complimentary materials solutions. In this effort methods both analytical and numerical have been developed and used for initial design of research rocket chamber-nozzle assemblies have been developed. These methods are general and can be used in future follow efforts.

The aerospike configuration or spike itself proved to be a significant heat transfer technical challenge. The predicted hot-fire durations with the best materials such as refractory composites such as carbon silicon carbide were found to be on the order of seconds. A first stage flight time for a typical rocket trying to put a payload in orbit is on the order of minutes. The un-cooled configurations as conceptually designed and analyzed simply do not work.

These results have shown cooling approaches which were considered early in this effort and dropped because of complexity must be reconsidered.

These findings partially motivated an external proposal opportunity to study cooling approaches to solve the problem discovered in this effort. A NASA Phase 1 STTR Proposal was written this past fall. This proposal was funded[1] and is now underway.

The heat-flux experiment that was developed will be hot-fired to calibrate these findings and be used in future development.

---

[1] NASA-STTR Phase 1 proposal "A Reusable, Oxidizer-Cooled, Hybrid Aerospike Rocket Motor for Flight Test, - 01/16/09. Principal Investigator: W.R. Murray, co-Investigators: T.W. Carpenter, J.D. Mello, and P. Lemieux. $100,000: [Rolling Hills Research Corp., $42,500; Cal Poly, $57,500].

The work on this project made possible by the Department of the Navy, Office of Naval Research, under Award # N00014-06-1-1111. This support enabled two students to learn an incredible amount and also resulted in findings that are being used now in follow-on work.

Attachment 1. - Preliminary Design Study of an Un-cooled Annular Aerospike Nozzle

A Thesis

presented to

the Faculty of the College of Engineering

California Polytechnic State University, San Luis Obispo

In Partial Fulfillment

of the Requirements for the Degree

Master of Science in Mechanical Engineering

by

William R. Browning

June 2007

ABSTRACT

Title: Preliminary Design Study of an Un-cooled Annular Aerospike Nozzle

Author: William R. Browning

Applied research has been conducted in the form of a preliminary design study to assess the feasibility of an un-cooled labscale annular aerospike nozzle. This study supports ongoing research at the Cal Poly Hybrid Rocket Motor Facility to explore this nozzle geometry as an alternative to a conventional nozzle. A by-product of this design study was the creation of three design tools that can be used as a foundation for future design studies. These design tools include the development of a CFD model that predicts gas-side surface film coefficients and a method of applying these coefficients to a two part FEA model. In the first part of the FEA model, these coefficients are applied to a transient heat conduction model to create a temperature field throughout the nozzle components. This temperature field is then mapped into a thermal structural model to predict thermal and mechanical stresses within each component. These tools were used to study the application of different high temperature materials to various components in the nozzle. Materials used for the spike component that is positioned at the center of the nozzle were given special attention as it represents the largest design hurdle. The results of this preliminary design study estimate a maximum operating time of approximately two seconds with regard to the first instance of failure predicted by the FEA model for each component. These results have serious implications about the feasibility of an un-cooled nozzle design for the aerospike configuration and suggest that cooling mechanisms be utilized in the spike.

ACKNOWLEDGEMENTS

This thesis is dedicated to my wife: Thank you for moving away from your school so that I could go to mine. Your support has been a driving force throughout this effort. Your love, patience and understanding mean everything to me. Your knowledge of Fortran helped, too.

I would like to extend my deepest gratitude to the members of the Mechanical Engineering department at California State Polytechnic University in San Luis Obispo for always having your doors open. In particular, I would like to give a special thanks to Professor Joe Mello and Professor Kim Shollenberger. Dr. Shollenberger always made an effort to answer my "quick questions" and provided endless support for the CFD study in this thesis. Dr. Mello's experience and expertise have been invaluable in supporting my ambitions to do "significant work".

CHAPTER 1

Introduction

A recent interest in commercial space exploration is evident by endeavors like the X-prize competition and accessible human space travel offered by Spaceship One. This interest has prompted research at the Cal Poly Hybrid Rocket Motor Facility to revisit a rocket nozzle technology that began development in the 1960's by NASA [1]. Recent interest in annular aerospike rocket nozzles for single stage rocket motors has primarily been because of its potential to be throttled. In a conventional converging-diverging nozzle active throttling is not possible and as a result it can not be optimized for a range of altitudes. The potential for throttling an aerospike nozzle lies with the spike itself. The idea is that the spike can be translated fore and aft along the axial coordinate thereby making the size of the annulus at the throat larger or smaller. Changing this annulus has a significant impact on the operating characteristics of the rocket motor, characteristics that could be tuned for different altitudes. In application, the ability to throttle is advantageous for sending small satellites into lower earth orbit.

Current research efforts are ongoing at the Cal Poly Hybrid Rocket Motor Facility that seek to understand and design mechanisms that could be used to actualize this potential for throttling and also understand thrust vectoring [1]. In addition, hot-fire testing using a conventional converging-diverging nozzle has been initiated to study different fuel grains and oxidizer injection configurations. Current labscale testing uses liquid nitrous-oxide as an oxidizer and either solid hydroxyl terminated polybutadiene (HTPB) or Plexiglas for the fuel grain. This research is directed at creating a labscale

annular aerospike nozzle to further investigate the design and implementation of control mechanisms for throttling and thrust vectoring.

A rendering of the aerospike nozzle assembled onto the current testing appartus is shown in Figure 1.1. Unlike a conventional nozzle there is no hardware that defines the supersonic region except, for the tip of the spike. The contour of the spike tip is dependent on the properties of the combusted gas and the desired operating conditions of the rocket motor for ideal expansion of the exhaust plume [2]. The diameter of the nozzle liner at the throat is also dependent on these characteristics. The nozzle liner is seated against the spike web and retained by the nozzle outer. The spike outer is assembled to the testing apparatus through a Marman (V-band) clamp and holds the nozzle liner and spike web in place.

The retention of the spike presents a challenge for assembling some of the components of the apparatus as a whole. The bulbous portion of the spike prevents regular assembly and disassembly of the inner fuel grain and spike web. There are two attachment strategies that have been proposed for axially retaining the spike. In one approach the spike is allowed to separate at some point within the rocket motor, while the other approach puts the point of separation on the outside, at the bottom of the testing apparatus. The latter approach would not be easy to implement with the current testing apparatus because there is little room to access its lower region. The first approach offers a greater element of serviceability by offering a more accessible point of separation. However, placing a joint inside the rocket motor creates sealing issues that could contribute to failure of the spike. A suggested approach to this attachment is made at the conclusion of this design study. The spike web that lies at the inlet of the nozzle is a

spoked structure intended to provide lateral stability for the spike so that it stays eoneentrie with the nozzle liner. This has been neeessitated by the potential for defleetion of the spike due to bending resulting from the length required to span the inner fuel grain. A perspeetive rendering of the nozzle liner, spike and spike web is given in Figure 1.2. A set of working drawings with dimensions can be found in appendix L.

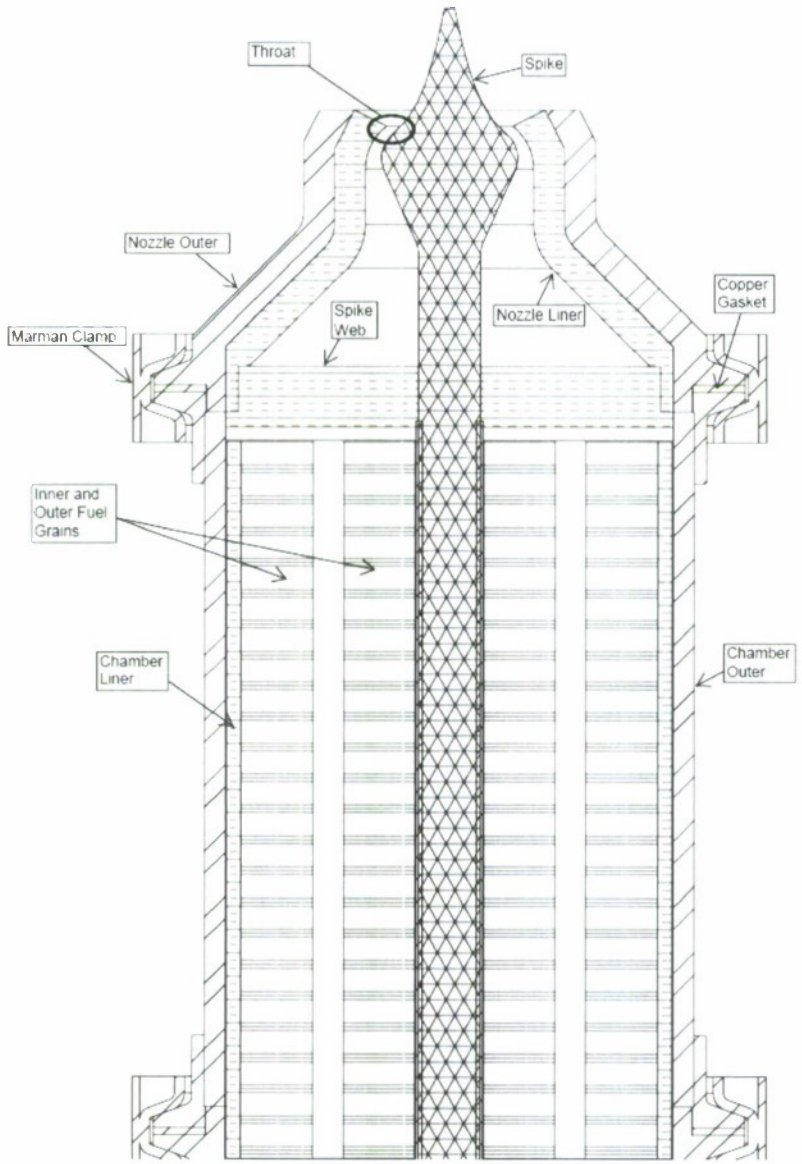


Figure 1.1 Rendering of the annular aerospike nozzle assembled onto the eurrent testing apparatus [1].

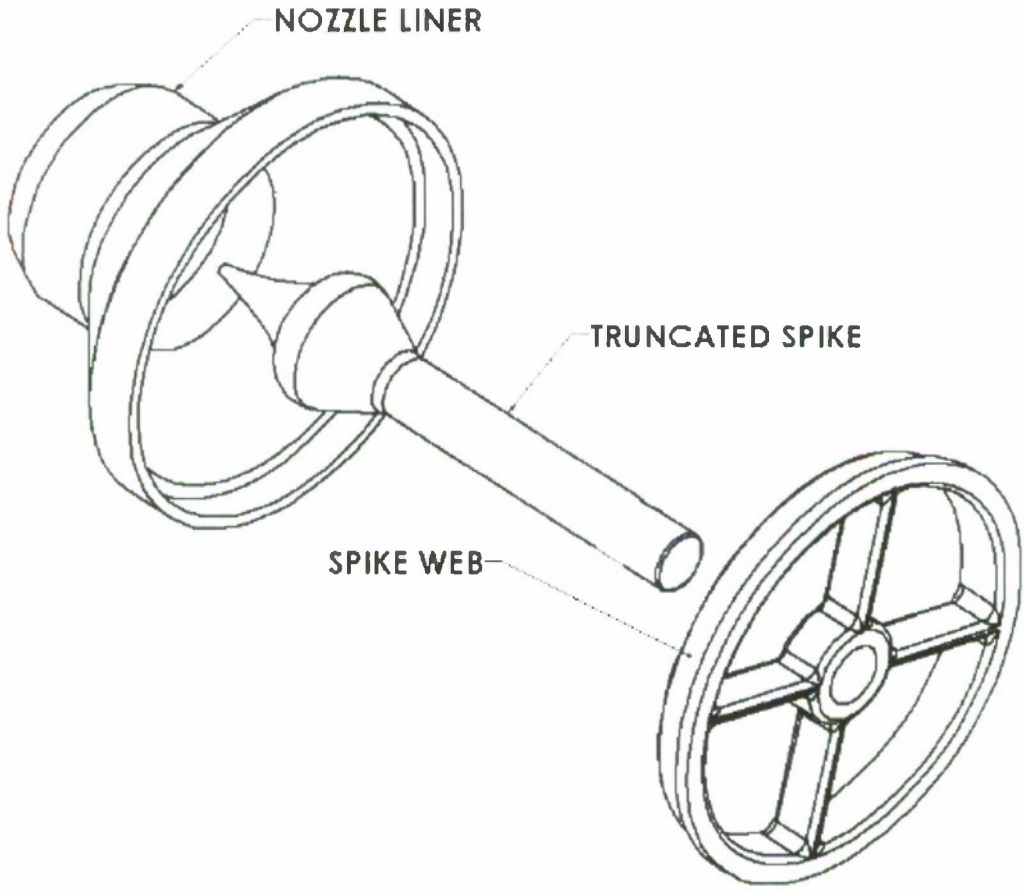


Figure 1.2 Exploded view of nozzle liner, spike, and spike web.

During a hot-fire seenario the spike will be surrounded by eombusted gases on all sides that are at elevated temperatures and pressures. Coupled with the high velocities inherent in a rocket motor, the gas-side conditions for this rocket nozzle geometry yield a highly strained flow with dominate viseous heating effects at the nozzle wall and surface of the spike. The highest rate of heat transfer typically exists at the throat for a conventional nozzle and the same trend is expected for the aerospike geometry [3]. In this harsh environment the spike is of particular concern because there are very limited opportunities for the spike to transfer the heat passed to it by the combusted gases. This results in high thermal stresses generated by large temperature gradients. A successful design of the spike depends on the strategy employed to overcome these large rates of

heat transfer. In addition, serious attention needs to be paid to the material(s) used in spike construction to prevent failure during its use.

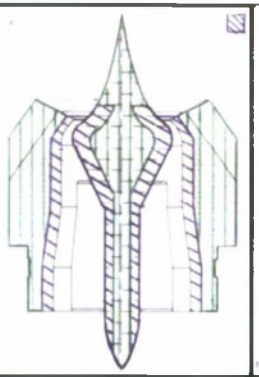
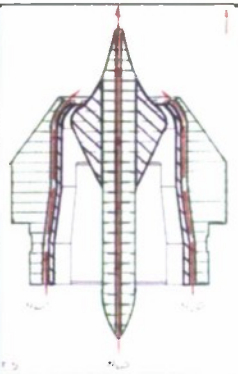
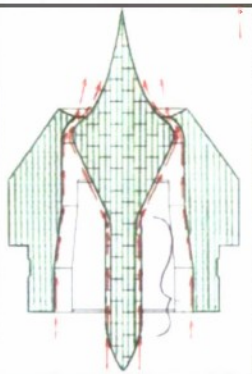
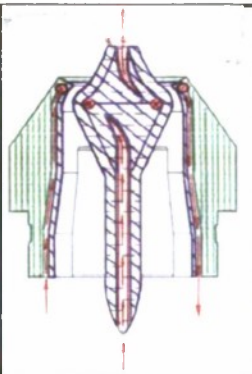
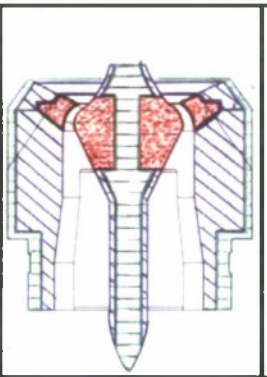
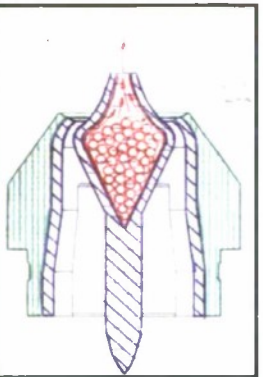
There are two groups known to have successfully conducted flight testing on rocket motors using an annular aerospike nozzle. Published results from these flight tests offer evidence of material behavior and damage for this nozzle configuration. In 2005 Trong T. Bui, *e. al.* flew the first known solid rocket motor equipped with this nozzle [4]. Two such rockets were launched along with a conventional nozzle version for comparison. Pressure-time traces from flight testing for these nozzles suggest an average chamber pressure just below 300 psi over approximately 7 seconds. Post-flight discussion mentions significant cracking of the graphite nozzle liner and spike. The cause of this damage is suspected to be related to thermal expansion [4]. In 2003, California State University at Long Beach (CSULB) flew two liquid rocket motors using both conventional and annular aerospike nozzles. This effort was preceded by a ground test in which an aerospike nozzle experienced catastrophic failure [5]. The spike that failed was constructed of solid graphite, and after less than a second of operation the spike experienced tensile failure at a point just upstream of the bulbous region. The design of this spike was revised for flight testing to include a centered titanium rod throughout most of the spike interior. This revision aided in the success of the flight testing. Published results suggest an average chamber pressure between 250 and 300 psi (peak pressure ~375 psi) over an approximate 7 second burn time [6]. Post-flight results also discuss significant damage of the nozzle liner and spike near the throat. In the revised CSULB design, the spike is seperated into two components that just aft of the throat. This was done to allow for the installation of the titanium rod through the center.

There seems to be evidenee of erosion at the mating surfaee between these two pieees. In addition, there are signifieant eraeks on the bulbous region of the spike and the nozzle liner near the exit suggesting that thermal expansion played a role in the damage of the CSULB aerospike nozzle during operation. It should be noted that the size of the testing apparatus has limited any potential design to a seale that is approximately two to three times smaller than the nozzles diseussed here.

As a first step towards flight ready annular aerospike nozzle hardware, a preliminary study has been eondueted to determine the best design approaeh. Special attention has been given to the spike, but without neglecting the other eomponents of the assembly. This study seeks to understand two aspects of the behavior of the nozzle assembly during a hot-fire seenario: the gas-side wall eonditions of the nozzle from the ehamber to the throat and the thermal and meehanieal stresses that arise in the material from the heat transferred to it by the flow of eombusted gases. The seope of this study has been narrowed to foeus only on a labseale version of the annular aerospike hardware with a stationary spike. However, lessons learned here will be applieable to future labseale and flight weight efforts. To initiate the design proeess several eoneepts were generated. The goal of this exereise was to determine a working design for this preliminary study. Six eoneepts were ereated that eonsider various eooling mechanisms and material eombinations. A visual summary of these eoneepts is given is Table 1.1.

The simplest of these designs are the *Heat Sink* and *Insulated* eoneepts. These designs would depend on the materials ability to withstand the eonditions of a hot-fire seenario without any aetive or passive eooling. The *Heat Sink* eoneept requires a thiek nozzle outer to provide a suffieient amount of mass. In this eoneept, a high temperature

Table 1.1 Early design concepts for the annular aerospike nozzle.

Concept Sketches						
Cooling Type	Heat Sink(minimal ablation)	Active (wasted N_2O)	Film	Active/ Regenerative(wasted N_2O through spike)	Insulated(minimal ablation)	Phase Change(closed or open tip)

composite, like earbon-carbon(C/C) or earbon-silicon carbide(C/SiC) would be used to eonstruet the spike and nozzle liner. An alternative material for this coneept could be graphite or an ablative, like earbon-phenolie. The *Insulated* concept would locally employ a high temperature eomposite near the throat and rely on graphite or carbon-phenolie for the remainder of the nozzle interior surfaces. This eoncept would eut down on eost by limiting the use the C/SiC, but eould have potential sealing issues at the interface. In both of these coneepts a metallie eore for the spike is envisioned for strength and serviceability.

The *Active*, *Film*, and *Active/Regenerative* coneepts utilize the nitrous-oxide (oxidizer) as a working fluid, taking advantage of its extremely low supply temperature. These coneepts offer the potential to be less aggressive with the material choiee for the nozzle interior beeause of lower gas-side wall temperatures. The drawback to these coneepts is that oxidizer is wasted in the eooling process. Wasting oxidizer is not neeessarily a coneern in a lab-seale environment, but it would impose serious performanee limitations for a flight weight seenario. In addition, the incorporation of active eooling meehanisms like regenerative or film eooling would add an element of eomplieation and eould be eostly. However, these options have been eonsidered because the need for aetive eooling may be unavoidable.

Lastly, the *Phase Change* coneept is interesting beeause it relies on the high latent heats of substances like sodium and silver for heat absorption. The working material of this concept would transition from a solid phase to a liquid or even a gas phase(sublimation) in order to remove heat from the system. Another variation of this design would elose off the spike tip so that the eooling mechanism is more like a heat

pipe. This design raises concerns about the chemical compatibility of the working material with a combustion environment, and would also add another level of complication to the design. Substances like sodium are highly reactive with H_2O which is found in the products of combustion.

The results of this exercise narrowed the set of concepts down to two designs. Concepts involving either active or passive cooling were abandoned because they are too complex for the scope of this project. However, it is reasonable to expect that future efforts, building upon this one, would include some form of active cooling. With the two remaining concepts relying solely on the ability of the hardware to withstand the harsh environment of a hybrid rocket nozzle, the task now becomes selecting materials than can perform in this environment.

To effectively study the best material combinations that are to be employed in an un-cooled annular aerospike nozzle, gas-side conditions must first be understood. Understanding the thermal exchange between the wall of a rocket nozzle and the combusted gases flowing through it has been the subject of recent research. There are a number of semi-empirical correlations that can be employed to predict local Nusselt numbers as a function of the axial coordinate for a conventional nozzle. Unfortunately, there are no studies known to exist for annular aerospike nozzles. An effort has been made with this research to provide a practical CFD model for predicting gas-side Nusselt numbers for annular aerospike nozzle geometries. This investigation will not include the effects of radiative heat transfer. This exclusion is founded on the assumption that radiation does not contribute significantly to the overall heat transfer at the throat, where it is highest [7].

Commercial CFD code Fluent 6.2.16 and meshing preprocessor Gambit 2.3.16 were used exclusively to model the turbulent flow through the aerospike nozzle. This CFD model investigates different strategies for turbulence model selection, meshing and data analysis. Verification for this CFD model is based on a conventional converging-diverging nozzle geometry and a survey of relevant semi-empirical correlations for predicting Nusselt numbers along the wall of this nozzle type. Results from the CFD model will be directly applied to an FEA model to analyze material behavior.

Thermal and mechanical stresses that are generated in the components of the annular aerospike nozzle will be analyzed using a two-part FEA model. Commercial FEA code Abaqus 6.6-1 is used to model the thermal loading of the annular aerospike. The thermal loading is simulated by applying gas-side film coefficients (representing convection only) to the nozzle liner and spike surfaces in a transient heat conduction model. The resulting temperatures throughout each component are mapped into a thermal structural model that predicts stresses. The loading of the Marman clamp is also considered. Within these models different materials are applied for evaluation in the overall design. Results for these models are compared and discussed. Concluding remarks include a suggestion regarding the best design for an un-cooled annular aerospike nozzle and a proposed time limit for operation. In addition, the restriction of the current design goals that limit the nozzle configuration to one that is un-cooled are reexamined.

CHAPTER 2

Material Selection

There are many factors to consider when choosing materials for a rocket motor design. Erosion, resistance to oxidation and strength are among the more important considerations for material selection. Unfortunately, cost and availability have a significant impact as well. The best marriage of high temperature materials will yield the longest sustained burn times and highest durability. In the aerospike nozzle configuration the spike will experience the harshest conditions and therefore presents the greatest challenge for material selection. Graphite, high temperature composites and various alloys have been surveyed for use in this design study for all of the components in the nozzle assembly.

Alloys can be ideal for rocket applications because in most cases they are cost-effective, durable, have good resistance to oxidation, have high strength, have good workability and are easier to analyze. However, some drawbacks are that alloys typically have operational temperatures that quickly can be exceeded in a rocket motor. In addition, weight becomes a major issue for flight design. Four different alloys were used in this design study. The application of these materials to the nozzle has been limited to avoid direct exposure with combusted gases. Stainless steel 304L is the most common among these four and will be applied to nozzle outer. This alloy is currently being employed in the same capacity for the conventional nozzle testing and performs well. However, an adjustment in the material thickness for this component may be required for the aerospike configuration. The other three alloys used in this study have higher operational temperatures than the stainless and are applied to the core of the spike and the

spike tip. Inconel 625 and Haynes 188 have similar material properties documented up to 1144 K (1600 °F) and 1255 K (1800 °F), respectively [8]. The fourth material is Columbium FS85, which has comparatively higher operational temperature with properties documented up to 1700 K (2600 °F) [8]. However, Columbium FS85 is not as cost-effective due to limited availability.

Graphite is a common material used at the throat of conventional rocket nozzles [3]. This material is typically employed as an insulator to protect metals from direct exposure to high temperatures. Graphite has a tensile strength that is approximately one third of its compressive strength, making it inherently poor in tension [9]. Another drawback of graphite is that while it has the ability to withstand high temperatures it has a poor resistance to oxidation. This poses a serious concern for a hybrid rocket because of the oxidizer. Industry professional Dr. Robert Shinavski at *Hypertherm* suggested applying a silicon carbide (SiC) ceramic coating to the surfaces of components made of graphite as away to increase the resistance to oxidation [10]. An ideal grade of graphite for this coating is aptly named SiC-6 and is used for the nozzle liner and outer surface of the spike for the un-cooled nozzle design.

Composite materials represent the forefront of rocket motor material technology. These materials have a higher cost and require more complicated analysis due to orthotropic properties. The benefit of using composites is a greater versatility in design. A component made from composite materials can be "tuned" for its application by controlling the orientation of the fiber direction. In doing so, the component design can be optimized for weight. There are a variety of different types of composite materials that can be incorporated into rocket motors that serve different purposes. A common

ablative material is carbon-phenolic, which is typically applied at the throat or used as an insulator [3]. However, since a design goal of this study is to minimize erosion, the application of an ablative would not be ideal. Carbon-carbon(C/C) composites offer better erosion behavior and have a higher strength than the carbon-phenolic.

Bunker and Prince compared the erosion rates of C/C to that of carbon-phenolic and silica-phenolic [11]. In this study ablation rates are calculated for solid and hybrid rocket motors. The chamber pressure for which these rates were calculated is 4.14 MPa (600 psi) with a chamber temperature of 3383.7 K (6090.7 °R) for the solid rocket motor and 3663.7 K (6594.7 °R) for the hybrid rocket motor. Their results show that C/C has about half the rate of erosion of carbon-phenolic and a quarter of the rate for silica-phenolic if using a hybrid motor. A C/C composite would be ideal for its low rate of erosion, but is cost prohibitive [12]. Also, C/C is subject to the same poor oxidation resistance as graphite and would likely have to be coated with a ceramic material, adding further cost.

Carbon fiber reinforced silicon-carbide (C/SiC) is proposed herein to be a more practical material for an un-cooled hybrid rocket motor. This high temperature composite boasts a greater resistance to oxidation than C/C by virtue of the ceramic matrix [13]. In addition, the C/SiC is about half the price of the C/C composite [12]. The use of C/SiC was further explored for application in the nozzle liner and the spike with the help of *Hypertherm*.

Input from Dr. Robert Shinavski has suggested that the *Insulated* concept could be modified so that the entire spike is made using 2.5D C/SiC [10]. It was determined that tooling and labor costs for making a prototype parts out of C/SiC make up the bulk

of the overall eosts. For this reason, it would be just as eost effective to construet the entire spike using C/SiC as opposed to a loeal applieation as represented in Table 1.1. Furthermore, the added strength would eliminate the need to use a metallic core in the spike to handle axial and thermal loads. This modifieation is ineluded in the design study along with the original approaeh that uses a metallic core.

The fabrieation of sueh a design would start with a eomposite preform of C/SiC, whieh is essentially a rectangular briek. In the preform, fibers are oriented in three orthogonal material direetions, two of whieh have the same properties. However, this *does not* mean that the material exhibits in-plane quasi-isotropy [14]. The brick would then be maehined using special diamond eoated tooling to ereate the shape of the spike. There is one drawbaek to this approaeh that affeets the analysis. The spike has an axisymmetrie geometry, but if this eomposite is used then it will likely exhibit mechanieal behavior that is not axisymmetric. For example, if the spike is uniformly heated it may expand more in one prineipal material direetion than another. This would cause the throat to deviate from representing two eoneentric annuli and may affeet engine performanee. In addition, a variation in eoeffieients of thermal expansion speeifie for a given material direetion eould generate large stresses within the material, even with uniform heating.

Dr. Shinavski also suggested that the application of C/SiC for the nozzle liner would not be eost-effieient [10]. Conventional nozzles using graphite inserts that are currently being tested have displayed reasonable performanee and better behavior is expeeted of the nozzle liner in the aerospike eonfiguration if eoated with silieon-earbide. For this reason any effort to design a nozzle liner using C/SiC was abandoned.

The main drawbaek to using C/C or C/SiC is a laek of material properties. Material properties for eomposites are diffieult to measure and are often proprietary or sensitive with regards to national defense. Only limited information for the 2.5D C/SiC materials properties was available from *Hypertherm* for these reasons [10]. For the analysis of spike eoneepts using this material, theoretieal estimates for the orthotropic material stiffness matrix were used as surrogates [15]. Surrogate properties for specific heat, thermal eonduetivity, density and eoeffieients thermal expansion were obtained from a material data sheet for a C/SiC laminate available through General Electric [16]. Other assumptions were made regarding the applieation of these material properties and are diseussed in later ehapters.

CHAPTER 3

Solver Specifications and Boundary Conditions

The turbulent flow of hot gases through a rocket nozzle is complex in nature and difficult to analyze. The combusted produets react throughout the nozzle changing the properties of the fluid. The heating of the nozzle wall changes the properties across the boundary layer and can have an affect on local film coefficients. In an effort to simplify the prediction of gas-side film coefficients a series of assumptions have been made for both nozzles. First, the scope of the study is narrowed only to consider steady state conditions. The existence of high Mach number flow requires the assumption of compressibility. Furthermore, choked conditions are prescribed at the throat. To avoid any consideration of heat generation in the nozzle the produets entering it will be assumed fully combusted and the chemical composition frozen. Lastly, the turbulent flow is assumed to be fully developed. These assumptions are implemented in Fluent through the solver parameters, fluid property specifications and boundary conditions.

Fluent 6.2.16 is a pressure-based computational fluid dynamics solver. For this code the governing equations are discretized using a finite-volume method. For a generic property Φ the conservation equation [17]:

$$\int_A \rho\phi\vec{V}\cdot\vec{n}\partial A = \int_A \Gamma_\phi \nabla\phi\cdot\vec{n}\partial A + \int_V S_\phi \partial V \tag{3.1}$$

is discretized into the form:

$$\sum_f^{N_{faces}} \rho_f \phi_f \vec{V}_f \cdot \vec{n} A_f = \sum_f^{N_{faces}} \Gamma_\phi (\nabla\phi)_f \cdot \vec{n}_f A_f + S_\phi V \tag{3.2}$$

Where ρ is the density, V is the velocity vector, n is a normal veetor for the area, A is the area and Γ is the diffusion coefficient of Φ. The subscript f refers to the face of a 3-D

control volume. In the case of a 2-D geometry the faces become edges. The solver was set to "coupled" so that the governing equations are simultaneously solved for using fluid properties from the current solution iteration[17]. The SIMPLE algorithm is used for the pressure-based solver by default for the coupled selection. In addition, the solver was set for steady state flow, implicit time integaration and axisymmetric geometry.

There are four boundary types that need to be specified for both nozzle geometries in the CFD model. For the conventional nozzle (base case) there is an inlet, exit, nozzle wall and axis of symmetry. The aerospike geometry is similar except that instead of an axis of symmetry there is an inner wall that is defined by the shape of the spike. These boundary types are first defined in Gambit and then assigned as boundary conditions later in Fluent. The boundary types for both geometries are displayed in Figure 3.1.

Inlet Boundary Condition

At the inlet boundary a pressure and temperature are specified based on the conditions of the rocket motor combustion chamber. In addition to turbulent and fully developed the flow is assumed to be normal to the inlet plane. It is shown later that the assumption of fully developed turbulent flow is reasonable, but not necessarily accurate. Another assumption is that the combustion efficiency of the rocket motor is 100%. This means that the products entering the nozzle can be treated as fully combusted, further reducing the complexity of the model by eliminating any reactants in the nozzle flow. Finally, certain parameters have to be specified when using turbulence models in Fluent [17]. For the k-ε models used in this research, the turbulence intensity and turbulent viscosity ratio will be specified for the inlet and exit boundary conditions.

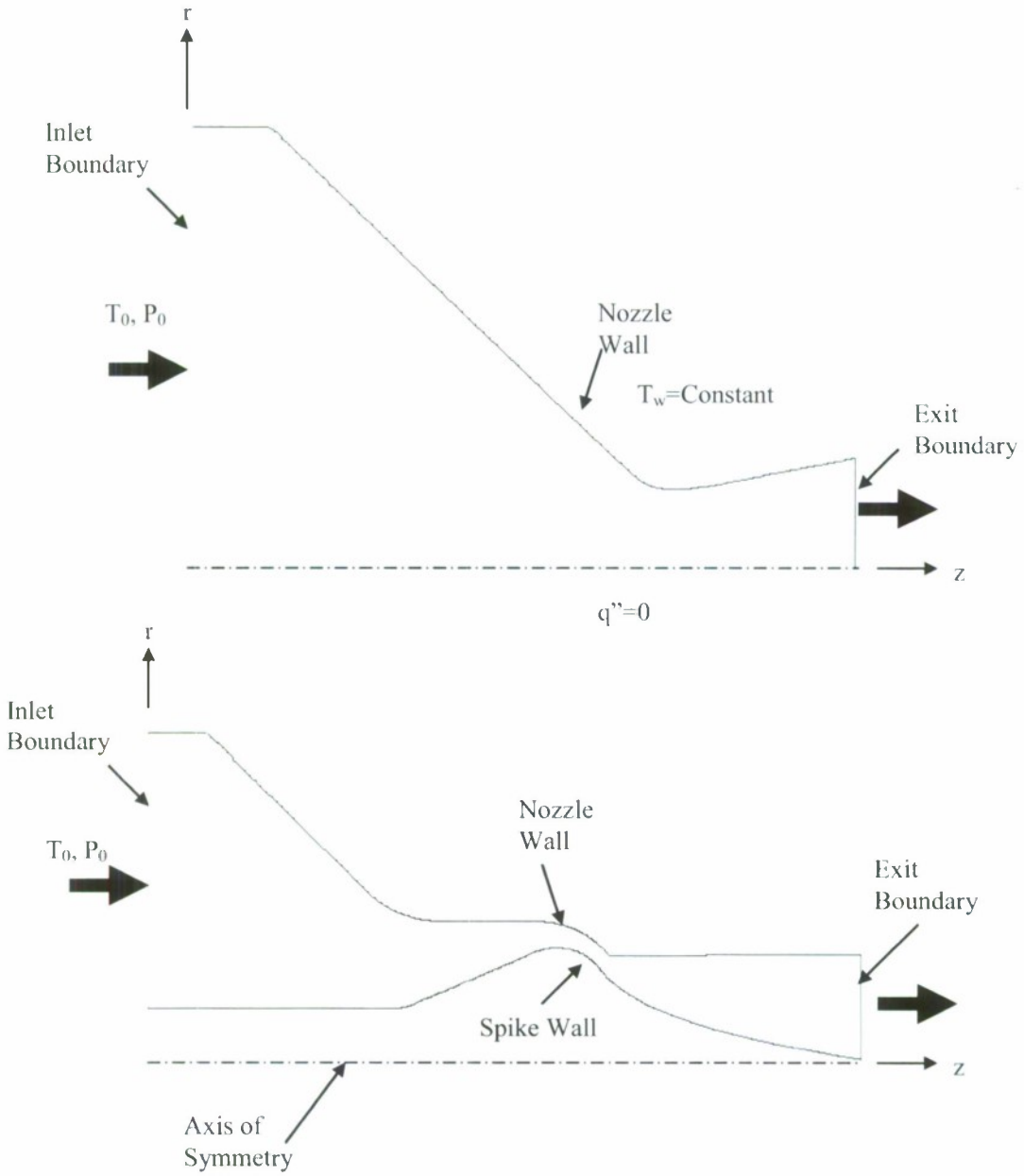


Figure 3.1 2-D axisymmetric geometries for both the base case and aerospike nozzles, respectively, are shown with boundaries labeled.

Exit Boundary Condition

The exit conditions for the base case model are not as involved because the exit plane exists at the end of the supersonic region and if choked flow is achieved ($M=1$ at the throat) then it does not matter what the exit conditions are because the upstream flow

will not be affected [3]. Based on this, the default exit conditions in Fluent could have been used. However, reasonable estimates for the exit temperature and pressure were made available during the process of defining the inlet conditions. It made sense to use these values instead of ambient conditions simply because of their availability.

The aerospike geometry is more complicated because the exit plane of the nozzle hardware and the throat are one in the same. This means that there is no hardware in the supersonic region making it difficult to apply exit conditions. After some trial and error it was determined that the best modeling approach was to create an artificial nozzle wall in the supersonic region that extends all the way to the end of the spike, making it the exit plane. The specific shape of the artificial nozzle wall is backed up by the idea that the optimum expansion of the exhaust plume is defined by a cylindrical boundary shape that is parallel to the axis of symmetry [2]. Because the shape of this supersonic region is similar to the predicted shape of the plume it is assumed that the flow is not bounded in such a way that is constricting. However, the shape should not affect the upstream conditions of the flow beyond the throat that are of specific interest to the CFD study. The exiting conditions at the end of the spike were based on a prescribed nozzle exit pressure ratio of 20.

Nozzle Wall

In order to capture the convective heat transfer coefficients along the wall, a surface temperature must be prescribed. Unfortunately, actual nozzle wall temperature data for the base case model would be difficult to measure and are not available. It is here that the fully developed turbulent flow assumption is beneficial. This assumption allows for a constant temperature to be assigned at the nozzle wall [18]. It should be

noted that since the nozzle wall for the aerospike model is extended beyond the throat to account for the artificial supersonic region it has the same boundary conditions imposed on the rest of the nozzle wall.

Lower Bound

In the base case study there is no wall defining the lower bound because at the center of a conventional nozzle is hot gas. But, because it is the center of the nozzle, an axis boundary is defined that essentially imposes an adiabatic wall condition at the lower bound of the model. The lower bound of the aerospike *does* have a wall that is defined by the spike geometry. For this reason a wall boundary type must be prescribed instead of an axis boundary. This wall will also have an adiabatic condition imposed upon it. This makes sense because even though there is material beyond the wall boundary there will not be any heat transfer in the radial direction for the steady state conditions.

CHAPTER 4

Turbulence Models

Fluent offers the *k-ε*, *k-ω* and *Spalart-Allmaras* turbulence models. The *k-ε* and *k-ω* turbulence models are two equation transport types, while the *Spalart-Allmaras* model is a single equation turbulence type. The *k-ε* model was used exclusively for this study because it has a proven track record for a wide variety of external and internal flows [19]. The use of this turbulence model assumes that the turbulent flow is fully developed. Two variations of the *k-ε* model called *Standard* and *RNG* were employed for this research.

Standard k-ε Turbulence Model

The *Standard k-ε* model is the least complicated of the variations and is a semi-empirically based two equation transport model. It has been used widely in industrial applications since 1972 when it was introduced by Launder and Spalding [17]. Two equations were derived by Launder and Spalding. One equation is for the turbulent kinetic energy *k* and the other is for the turbulent dissipation rate *ε* [17]:

$$\frac{\partial}{\partial t}(\rho k)+\frac{\partial}{\partial x_i}(\rho k u_i)=\frac{\partial}{\partial x_j}\left[\left(\mu+\frac{\mu_t}{\sigma_k}\right)\frac{\partial k}{\partial x_i}\right]+G_k+G_b-\rho\varepsilon-Y_M+S_k \tag{4.1}$$

$$\frac{\partial}{\partial t}(\rho \varepsilon)+\frac{\partial}{\partial x_i}(\rho \varepsilon u_i)=\frac{\partial}{\partial x_j}\left[\left(\mu+\frac{\mu_t}{\sigma_\varepsilon}\right)\frac{\partial \varepsilon}{\partial x_i}\right]+C_{1\varepsilon}\frac{\varepsilon}{k}(G_k+C_{3\varepsilon}G_b)-C_{2\varepsilon}\rho\frac{\varepsilon^2}{k}+S_\varepsilon \tag{4.2}$$

where,

$$\mu_t=\rho C_\mu \frac{k^2}{\varepsilon} \tag{4.3}$$

$$G_k=-\rho\overline{u_i'u_j'}\frac{\partial u_j}{\partial x_i} \tag{4.4}$$

$$G_b = -g_i \frac{\mu_t}{\rho \mathrm{Pr}_t} \frac{\partial \rho}{\partial x_i} \tag{4.5}$$

$$Y_M = 2\rho\varepsilon\left(\frac{k}{a^2}\right) \tag{4.6}$$

The model constants in Fluent are as follows: $C_{1\varepsilon}$=1.44, $C_{2\varepsilon}$=1.92, C_μ=0.09, σ_k=1.0, σ_ε=1.3.

On the right hand side, G_k represents the produetion of turbulent kinetic energy, G_b represents the generation of turbulcnce due to buoyaney, and Y_M is called the "dilatation dissipation", whieh is important for high Maeh number flows. The a^2 term in the equation for Y_M is the squarc of the speed of sound at a given speeifie heat ratio and temperature [17].

RNG k-ε Turbulence Model

The *RNG k-ε* model differs from the *Standard* variation in that it is derived using statistical technique instead of being semi-empirically based. In faet, *RNG* stands for the statistieal tcchnique ealled "renormalization group theory". The *RNG k-ε* model offers advantages over the *Standard k-ε* model for low-Reynolds number effects making it more accurate for a larger range of flows. One distinction is the use of a differential equation to solve for the effective turbulent viscosity (Equations 4.9 and 4.10). Integrating this equation offers better accuracy in handling low-Reynolds number and near-wall flows. The form of the two transport equations is very similar to the *Standard k-ε* model[17]:

$$\frac{\partial}{\partial t}(\rho k) + \frac{\partial}{\partial x_i}(\rho k u_i) = \frac{\partial}{\partial x_j}\left(\alpha_k \mu_{eff} \frac{\partial k}{\partial x_i}\right) + G_k + G_b - \rho\varepsilon - Y_M + S_k \tag{4.7}$$

$$\frac{\partial}{\partial t}(\rho \varepsilon) + \frac{\partial}{\partial x_i}(\rho \varepsilon u_i) = \frac{\partial}{\partial x_j}\left(\alpha_k \mu_{eff} \frac{\partial \varepsilon}{\partial x_i}\right) + C_{1\varepsilon}\frac{\varepsilon}{k}(G_k + C_{3\varepsilon}G_b) - C_{2\varepsilon}\rho\frac{\varepsilon^2}{k} - R_\varepsilon + S_\varepsilon \tag{4.8}$$

Where μ_{eff} is solved for through:

$$\hat{\nu} = \frac{\mu_{eff}}{\mu} \tag{4.9}$$

and,

$$d\left(\frac{\rho^2 k}{\sqrt{\varepsilon\mu}}\right) = 1.72\frac{\hat{\nu}}{\sqrt{\hat{\nu}^3 - 1 + C_\nu}}d\hat{\nu} \tag{4.10}$$

For high-Reynolds number flows Equation 4.10 reduees to Equation 4.3. In addition, it should be noted that when this reduetion takes plaee a different value of C_μ is used. For this k-ε Turbulenee Model C_μ is derived using *RNG* theory, resulting in a value of 0.0845 instead of 0.09. The *RNG* model constants are as follows: $C_{1\varepsilon}$=1.42, $C_{2\varepsilon}$=1.68, C_ν=100.

An extra term ealled R_ε shows up in the turbulent dissipation transport equation. The formulation for this term is eomplex and will not be discussed in detail. However, its affect on the turbulent dissipation rate transport equation can be summarized. This term aligns the *RNG* model in good agreement with the *Standard* model when the flow is moderately strained. However, this agreement lessens as the strain rate inereases. In the end the *RNG* model prediets lower turbulent viseosities for highly strained flows eompared to the *Standard* model.

CHAPTER 5

Convergence Schemes

When the governing equations are discretized and applied to a numerical model the solution has to be converged upon iteratively. After each iteration the properties are updated in each cell of the mesh. The values that are used in the discretized form of the governing equations are taken at the faces of these cells. For an arbitrary fluid property Φ that characterizes an entire cell, the value at the face between two cells is Φ_f. The face values depend on the cell values, but there is one value upstream and one downstream of the face being examined that need to be considered. To simply take an average of Φ would not be physically accurate. It turns out that the value of Φ_f has a greater dependency on the upstream cell value in the flow than the downstream [19]. Various convergence schemes have been designed to deal with weighting this dependency. There are 5 convergence schemes available in Fluent for the turbulence kinetic energy and dissipation rate. Some are more accurate than others and can offer performance advantages for certain types of meshes.

First-Order Upwind

The simplest of these schemes is the first-order upwind scheme, sometimes called upwind differencing. In this convergence scheme the face values are set equal the upstream cell values and have no dependency on the other cell[17]:

$$\phi_f = \phi_{up} \tag{5.1}$$

This scheme provides the least accuracy but can still be useful in starting a simulation for a large mesh. When dealing with large mesh sizes running a more accurate convergence scheme can be computationally expensive. To overcome this, the simulation can be initiated using the first-order upwinding scheme. Once the model converges to a crude result, the model can be continued using a more complicated convergence scheme that will eventually converge to a more desirable result [20]. This strategy was not necessary for the present study because the mesh was small enough that a more accurate convergence scheme was used from the beginning.

Second-Order Upwind

The second-order upwind scheme offers better accuracy than the first-order and is relatively efficient. This scheme offers a dependency on both cells adjacent to the face in between. The relationship essentially takes the first-order scheme and adds another term from a Taylor series expansion[17]:

$$\phi_f = \phi_{up} + \nabla\phi_{up} \cdot \vec{r} \tag{5.2}$$

The second term represents the gradient of Φ_{up} across the cell and the direction vector r captures the distance from the upstream cell centroid to the centroid of the face being studied. A 1-D representation of the second-order upwinding scheme looks like[19]:

$$\phi_f = \phi_{up} + \left.\frac{\partial\phi}{\partial x}\right|_{up} \partial x \tag{5.3}$$

Quick

An adaptation on the second-order upwinding scheme called Quick adds Equation 5.2 with a centrally interpolated value of Φ from neighboring cells. This centrally interpolated value reduces to a simple average if all the cell widths are equal. In this

form the contribution from the second-order upwind scheme and the average value is controlled by a weighting factor θ. The Quick scheme still depends mostly on the second-order upwind portion. However, the inclusion of the central interpolation scheme makes the Quick scheme more accurate for structured meshes where grids are aligned with the direction of fluid flow. Because of these performance advantages, this scheme was used for the present research.

CHAPTER 6

Base Case Study

The geometry of the annular aerospike nozzle interior is more complex than a conventional converging-diverging nozzle and there is no guarantee that the numerical methods used by Fluent to solve for the convective heat transfer coefficients will be accurate, although it is expected. Furthermore, there is no way of initially knowing which k-ε turbulence model set up works best for the operating conditions of the nozzle. For this reason a base case CFD study was conducted to determine this set up and to verify Fluent's ability to predict reasonable results. To effectively conduct such a study for the aerospike geometry requires the use of experimental and/or theoretical results that can be used as a benchmark. Unfortunately, no relevant results were found for this research. Instead, the base case study was modified to use the more traditional converging-diverging nozzle. This turned out to be a convenient modification because experimental data was readily available for a conventional nozzle that has been fired on the same apparatus intended for the testing of the aerospike nozzle [1, 21]. Basic information used for this base case study came from one such test firing labeled "*Run #20*" and can be found in Appendix A.

The convective heat transfer coefficients along the wall of the conventional nozzle could not be directly computed from the data available for "*Run #20*". Instead, these coefficients were approximated by first determining the combusted gas properties throughout the nozzle and then applying those properties to semi-empirical correlations for a local Nusselt number. For the latter part of this two-phase process a brief literature

survey was eondueted to ehoose a set of eorrelations that best approximates the heat transfer inside a eonventional roeket nozzle.

Survey of Semi-Empirical Correlations

There are a number of semi-empirieal correlations available that can be used to approximate the Nusselt number, and essentially the heat transfer eoeffieient, at any axial position within the fully developed region of turbulent flow through a eireular pipe or through a nozzle. The inelusion of eorrelations developed for pipe flow in this survey is predicated on the idea that their applieation ean be extended to turbulent flow through a eonventional converging-diverging nozzle by ehanging the restrietion of a eonstant diameter to one that varies with the axial eoordinate. In any ease, most correlations developed for pipe and nozzle flow share a eommon form based on the Chilton-Colburn analogy. The Chilton-Colburn analogy defines the Nusselt and Stanton number in terms of the frietion eoeffieient, C_f, as [22]:

$$\frac{C_f}{2} = \frac{f}{8} = St_D \cdot \mathrm{Pr}^{2/3} = \frac{Nu_D}{\mathrm{Re}_D \cdot \mathrm{Pr}^{1/3}} \qquad 0.6 < \mathrm{Pr} < 60 \qquad (6.1)$$

Where,

$$\text{Nusselt number(internal flow):}\ Nu_D = \frac{h \cdot D}{k} \qquad (6.2)$$

$$\text{Reynolds number(internal flow):}\ \mathrm{Re}_D = \frac{\rho \cdot u \cdot D}{\mu} \qquad (6.3)$$

$$\text{Prandtl number:}\quad \mathrm{Pr} = \frac{\mu \cdot C_p}{k} \qquad (6.4)$$

$$\text{Stanton number(internal flow):}\ St_D = \frac{Nu_D}{\mathrm{Re}_D\,\mathrm{Pr}} \qquad (6.5)$$

Also, f is the friction factor and can approximated for turbulent flow using the following correlation developed by Petukhov [22]:

$$f = (0.790 \cdot \ln \mathrm{Re}_D - 1.64)^{-2} \qquad 3000 \le \mathrm{Re}_D \le 5 \times 10^6 \tag{6.6}$$

The Colburn equation is a classic semi-empirical correlation and can be solved for by combining (6.1) and (6.6):

$$Nu_D = 0.023 \cdot \mathrm{Re}_D^{4/5} \cdot \mathrm{Pr}^{1/3} \tag{6.7}$$

The form of this equation can be generalized as:

$$Nu_D = K \cdot \mathrm{Re}_D^{m} \cdot \mathrm{Pr}^{n} \tag{6.8}$$

This form has been employed to predict Nusselt numbers in a variety of turbulent internal flows. Throughout most adaptations the leading coefficient K and Prandtl number exponent n vary slightly from the Colburn equation and such variations are usually founded in experiment. In 1930, F.W. Dittus and L.M.K. Boelter came up with two variations for pipe flow that are distinguished for cooling ($T_w<T_m$) and heating ($T_w>T_m$) by n [22]:

$$Nu_D = 0.023 \cdot \mathrm{Re}_D^{4/5} \cdot \mathrm{Pr}^{n} \tag{6.9}$$

where

$$n=0.3 \text{ for } T_w<T_m \tag{6.10}$$

$$n=0.4 \text{ for } T_w>T_m$$

Note that T_w is the wall temperature and T_m is the mean temperature of the fluid at a specific axial location. The Dittus-Boelter equations are perhaps most widely used for pipe flow, but some correlations have been developed specifically for conventional rocket

motor nozzles. S. Greenfield developed an equation for a liquid rocket motor having 4905 N (1103 lbf) of thrust and a chamber pressure of 1.962 MPa (284.56 psi) [23]:

$$Nu_D = 0.024 \cdot \mathrm{Re}_D^{4/5} \cdot \mathrm{Pr} \tag{6.11}$$

Experiments conducted by W.S. Long resulted in a form that deviates from Equation 6.8 by replacing the Prandtl number term with a ratio of the static temperature, or mean temperature, to the stagnation temperature [23]:

$$Nu_D = 0.024 \cdot \mathrm{Re}_D^{4/5} \cdot \frac{T_m}{T_0} \tag{6.12}$$

D.R. Bartz has done significant work in determining effective correlations for nozzle flow in a rocket motor. In a 1957 article published in *Jet Propulsion,* titled "A Simple Equation for Rapid Estimation of Rocket Nozzle Convective Heat Transfer Coefficients", he presented a direct estimation of the convective heat transfer coefficient throughout a nozzle that has shown good agreement with more complicated turbulent boundary layer methods [24]:

$$h_g = \left[\frac{0.026}{D_*^{0.2}} \cdot \left(\frac{\mu^{0.2} \cdot C_p}{\mathrm{Pr}^{0.6}} \right)_0 \cdot \left(\frac{P_c \cdot g}{C^*} \right)^{0.8} \right] \cdot \left(\frac{A_*}{A} \right)^{0.9} \cdot \beta \tag{6.13}$$

Where

$$\beta = \left(\frac{\rho_m}{\rho`} \right)^{0.8} \cdot \left(\frac{\mu_m}{\mu_0} \right)^{0.2} \tag{6.14}$$

$$C^* = \frac{P_c A_t}{\dot{m}} \tag{6.15}$$

The products of the terms within the brackets of Equation 6.13 are held constant while the area ratio and the β term are variable. β is defined as a correction factor that accounts for the variation in gas properties across the boundary layer which is an important

distinction that sets the Bartz formulation apart from previous examples. Within the β term the mean local density is divided by the local free stream value and the mean local dynamic viscosity is divided by the stagnation, or chamber condition, value. This correlation takes advantage of an assumption that the specific heat at constant pressure and the Prandtl number do not vary significantly along the axial coordinate and can therefore be assumed constant with regard to the stagnation temperature. The only caveat to this equation is that the correction factor requires some knowledge about the absolute viscosity and density of the gas at the wall so that an arithmetic mean value for these gas properties can be computed. Measuring gas-side properties near the wall is very difficult for high temperature, oxidizing environments. However, Bartz provides a set of curves that can be used to approximate β based on the specific heat ratio, area ratio and an approximate ratio of the wall temperature to stagnation temperature [24].

The range of applicability is broadest for the Dittus-Boelter equation since its uses extend beyond the realm rocket nozzles. This formulation has shown reasonable agreement for a range of Prandtl numbers from 0.7 to 160 and for Reynolds numbers above 10,000. Since the other three formulations were developed specifically for the flow of hot gases through a converging diverging nozzle they can only be assumed applicable to rocket motors. These correlations are presented here with expressed intention of providing a good first approximation for the prediction of film coefficients for the nozzle wall of the base case geometry. The hope is that as a collective their results will provide a reasonable and practical benchmark for developing the CFD model in the absence of relevant experimental data. To effectively apply these correlations requires a set of properties for the combusted gases throughout the nozzle.

Combustion Analysis of "Run#20"

The next step in the base case study is to determine a set of properties for the combusted gases throughout the nozzle that can be applied to the chosen set of correlations. These properties were not directly measured for the experimental data used and so they had to be calculated using combustion analysis software. The experimental data for "*Run #20*" was used in conjunction with a combustion analysis software called *Thermal Equilibrium Program*, or *TEP*, in order to provide a set of gas properties for the subsonic, throat and supersonic regions of the nozzle. Known operating conditions like chamber pressure (P_c) and the ratio of oxidizer to fuel (O/F) are typical inputs for *TEP*. In addition, the chemical formulas, enthalpies of formation, and densities for the propellant and oxidizer were specified. With these parameters in place combustion gas properties like local static temperature, density, specific heat at constant pressure, dynamic viscosity, Mach number, static pressure and gas thermal conductivity can be determined at any specified area ratio or pressure ratio.

In the case of a simple converging-diverging nozzle the use of area ratios are convenient because the geometries for both the subsonic and supersonic regions of the nozzle are simple cones whose area ratios change linearly with axial position. Thirty-one axial positions called "stations" were defined along the axis of symmetry. Fifteen evenly spaced stations were set before and after the throat and 1 station is positioned at the throat itself. However, not all 31 stations are evenly spaced. Since the subsonic and supersonic regions are different lengths the spacing between both sets of 15 are different. Specific gas properties can then be determined for each station by inputting the corresponding

area ratios into *TEP* for the combustion analysis. A diagram of these stations is given in Figure 6.1.

Table 6.1 Chemical composition of gas corresponding to chamber conditions.

Base Case Chemical Composition(Frozen)								
Species	N_2	CO_2	O2	H_2O	NO	CO	OH	O
Mole Fraction	0.574807	0.116998	0.111421	0.10033	0.030348	0.028297	0.022767	0.009292
Species	H_2	H	HO_2	NO_2	HNO	H_2O_2	N	
Mole Fraction	0.003351	0.002259	0.000078	0.000044	0.000003	0.000002	0.000002	

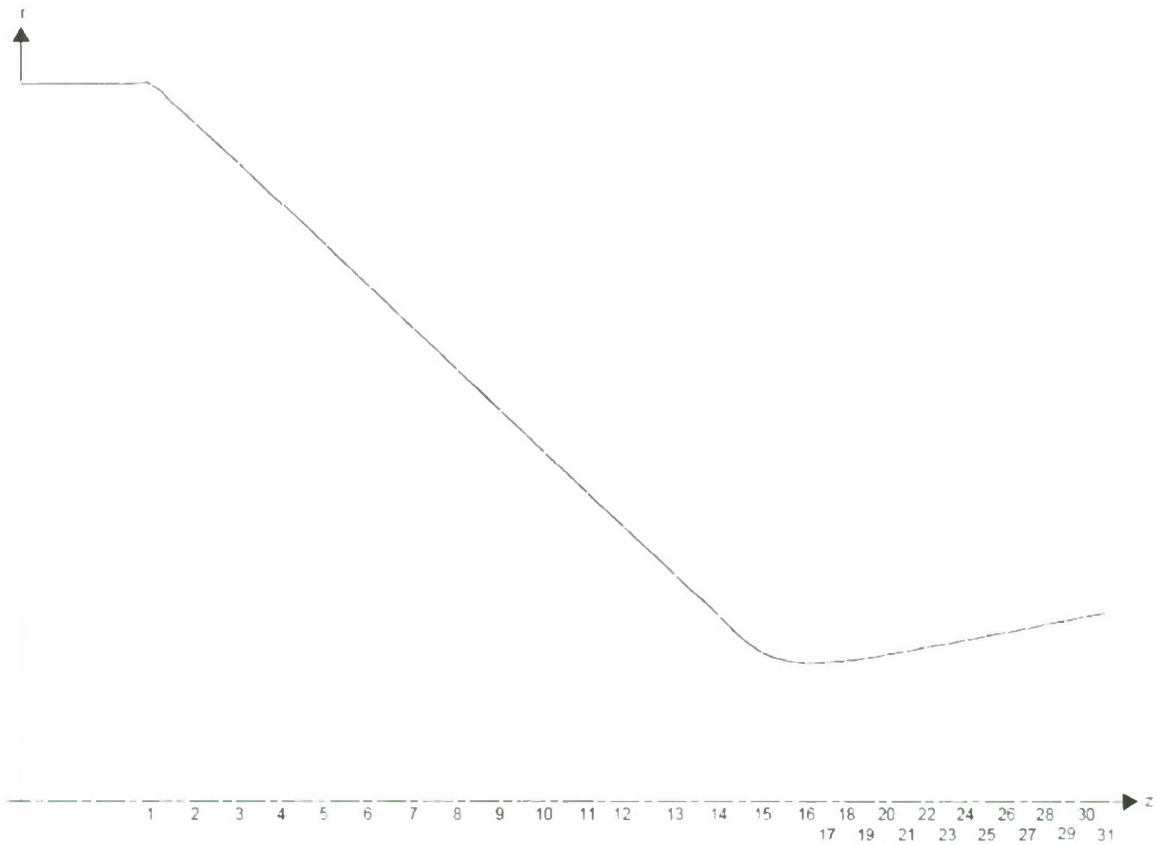


Figure 6.1 Converging diverging nozzle with 15 stations before and after the throat.

Gas properties at each station that are output by *TEP* come in two forms. One form assumes a frozen composition of the gas species based on the amount of each found in the combustion chamber and the other assumes an equilibrium composition. The species of the combusted gas change with area ratio for the equilibrium composition and typically have more water and carbon dioxide at the exit plane than at the inlet. Even though it seems that an equilibrium composition more accurately describes the combustion gas properties, working with both output sets can be quite practical. In fact, is has been shown that C^* efficiencies for equilibrium and frozen compositions tend to bracket actual measured efficiencies [21]. For this study gas properties resulting from a frozen composition have been used. This choice was made to compliment the way fluid properties are assigned in Fluent. A listing of the chemical composition used for the frozen assumption is given in Table 6.1.

When specifying the properties in Fluent for a fluid, there are input fields for density, dynamic viscosity, specific heat at constant pressure, thermal conductivity and molecular weight. For the CFD model the density was set to follow the ideal gas law because of the compressible flow assumption and the other parameters were held constant. By making the rest of the parameters constant a frozen composition condition is imposed on the CFD model with the exception of specific heat. However, holding the specific heat constant is still within reason, as it does not vary appreciably through the nozzle anyway. For the base case study the change in the specific heat was found to be approximately 4% across the nozzle from the inlet to exit plane. By modeling the fluid as having a frozen composition a greater level of consistency between the analytical and CFD models is achieved. A listing of gas properties at each station for the base case

study ean be found in Appendix B. A sample of these properties is given in Table 6.2 for the inlet, throat and exit locations of the nozzle. The gas properties for the ehamber are given for the inlet.

Table 6.2 Sample of gas properties assuming frozen eomposition used for the base case study.

Location	local diameter (m)	Area Ratio	Static Temp. (K)	rho (kg/m^3)	Mach #	viscosity (kg/s-m)	specific Heat (J/kg-K)	Gas thermal cond. (W/m-K)	Specific Heat Ratio	M_{wtmix} (kg/kg-mol)	Prandtl # (TEP)
Inlet	0.0809	27.38	2985	1.7460	0.0217	0.0001	1453.38	0.1705	1.2470	28.88	0.7129
Throat	0.0155	1.00	2657	1.0909	1.0000	0.0001	1439.98	0.1617	1.2454	29.13	0.7155
Exit	0.0212	1.87	1964	0.3332	1.9614	0.0001	1396.03	0.1393	1.2455	29.58	0.7197

Generating curves from Semi-Empirical Correlations

Gas properties gathered for eaeh station were applied to eaeh of the correlations to obtain a set of loeal Nusselt numbers. Exeept for the Bartz equation, eaeh eorrelation is already defined in terms of the Nusselt number. The Bartz equation had to be non-dimensionalized so that it could be effeetively eompared to other eorrelations. Computed Nusselt numbers were plotted against the non-dimensional axial position of eaeh station where the gas properties were evaluated and are displayed in Figure 6.2. In general the four eorrelations share the same trend, predieting the highest rate of heat transfer at the throat of the nozzle. These eurves will be used as the benehmark in verifying the CFD model set up for the base ease in Fluent.

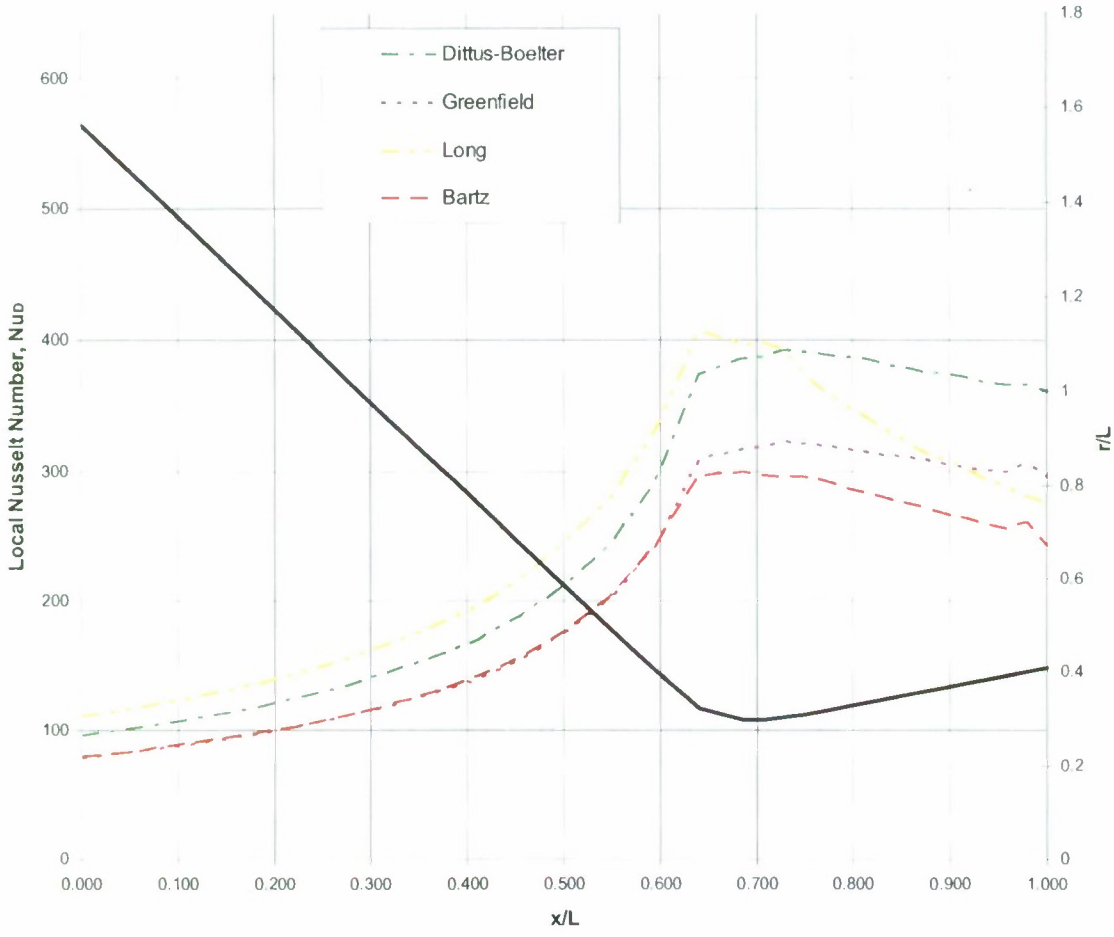


Figure 6.2 Local Nusselt numbers with respect to axial position. A curve for the nozzle wall is included for perspective.

CFD Model

In verifying Fluent, several different model setups were compared to the benchmark Nusselt number curves. The results from each model were analyzed by extracting local heat fluxes(q'') and mean temperature(T_m) data for each station. Local values for the convective heat transfer coefficient were obtained using Newton's Law of Cooling:

$$q'' = h_g \left(T_{aw} - T_w \right) \tag{6.16}$$

The T_{aw} term represents the adiabatic wall temperature is related to the mean temperature by [23]:

$$R = \frac{(T_{aw} - T_m)}{(T_0 - T_m)} \tag{6.17}$$

In Equation 6.17, R is called the recovery factor and is constant for fully developed turbulent internal flow. When the flow is fully developed, the shape of the temperature profile does not change with respect to axial position, although the mean temperature may increase or decrease in magnitude[18]. Experimental results show a strong relationship between this recovery factor and the Prantl number given by[23]:

$$R = \mathrm{Pr}^{0.33} \tag{6.18}$$

Combining equations 6.17 and 6.18 to solve for T_{aw} gives the following expression:

$$T_{aw} = \mathrm{Pr}^{0.33}(T_0 - T_m) + T_m \tag{6.19}$$

The local heat fluxes and adiabatic wall temperatures are applied to Newton's Law of Cooling along with the prescribed constant wall temperature to get local values for the heat transfer coefficients. These results are non-dimensionalized into local Nusselt numbers using the constant value of gas thermal conductivity prescribed for the fluid properties in the CFD model and local hydraulic diameters for each station. These Nusselt numbers are then compared to those from the semi-empirical correlations.

From these comparisons a decision was made regarding which k-ε turbulence model worked best. In addition, this study was used to determine if the fully developed turbulent flow assumption is reasonable as it pertains to the constant temperature boundary condition imposed along the nozzle wall. Another benefit of this base case

study was the development of a meshing strategy that can accurately capture the near wall effects of the flow through the nozzle.

The base case study can be summarized visually using the flow chart given in Figure 6.3. This flow chart seeks to put everything discussed in perspective by showing the hierarchy of processes and inputs. The first decision block determines if the near wall mesh is sufficient by analyzing results for the local non-dimensional turbulent boundary layer thickness, y^+. The meshing approach will be addressed in chapter 7, but suffice to say that if the answer is "*no*" then the mesh is reassessed. If the mesh is given a pass then another decision is evaluated that determines if results for a given model setup adequately agree with results for the semi-empirical correlations. This series of processes and inputs was also used in conducting the final CFD study for the aerospike configuration, but with some modifications.

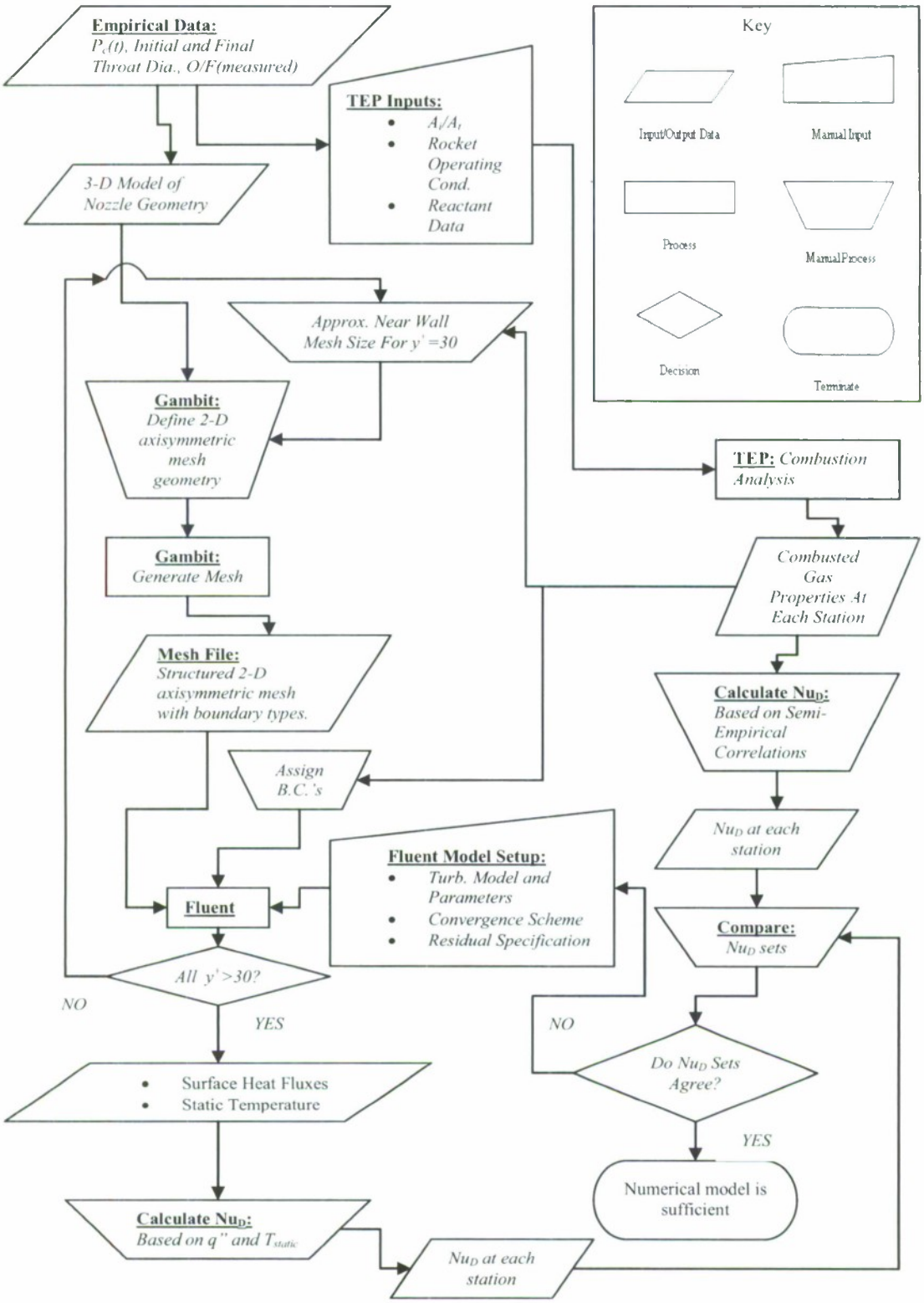


Figure 6.3 Process flowchart for the base case study with meshing refinement included.

CHAPTER 7

Meshing

The 2-D modeling capabilities of Gambit are not ideal for creating complicated 2-D geometries like that of the aerospike. For this reason CAD software was used to initially create both nozzle geometries. These models were then imported into Gambit using a STEP file format. Once imported the 3-D models are reduced to single 2-D profiles representing the axisymmetric nozzle flow fields shown in Figure 3.1.

Quadrilateral cells were used for the present research for simplicity and to cut down on computational expense [19]. Within Gambit there are two types of quadrilateral cell-based meshing schemes: structured and unstructured. Structured meshes are simpler and easier for computation. Using a structured mesh requires that the same number of nodes be used on adjacent boundaries. For example, if 10 nodes are applied at the inlet boundary then 10 nodes need to be applied at the exit if a structured mesh is applied. An unstructured mesh, on the other hand, allows for some deviation from this requirement, but can have numerical consequences. Unstructured meshes *do* come in handy when dealing with abrupt geometric changes where structured meshes would yield high aspect ratios. Although the aerospike geometry is complex enough to be created outside of Gambit, its smooth boundary curvatures made it a good candidate for a structured mesh. Also, since the base case geometry is simpler than the aerospike the application of a structured grid is even more ideal.

The structured meshing approach used in both the base case study and the aerospike study is separated into the two discussions: the free stream and near wall mesh. The free stream meshing scheme is fairly simple and straight forward, but the near wall

meshing strategy is a bit more involved. The handling of the near wall meshing scheme directly affects the accuracy of the local wall Nusselt numbers that are central to this CFD effort.

For fully developed turbulent flow the velocities and temperatures in the region near the boundary of flow can be described according to the universal law of the wall. According to this law, the turbulent boundary layer is defined by 3 layers that are called the viscous sublayer, buffer region, and fully turbulent region. The viscous sublayer is closest to the wall and it is here that the viscous dissipation dominates the transfer of momentum. A little further away from the wall, in the fully developed region, the turbulent dissipation becomes significant. In between these regions exists the buffer region where both dissipation terms dominate [25].

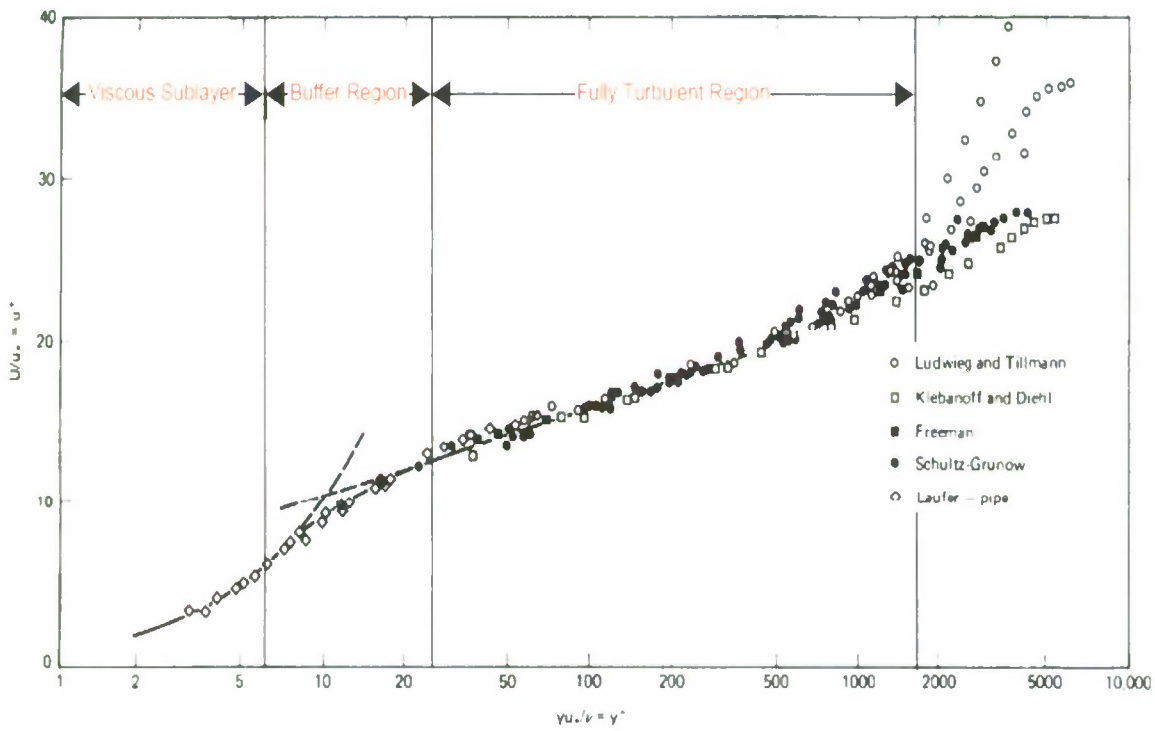


Figure 7.1 Typical universal law of the wall plot for a smooth surface with 3 layers shown [18, 26].

In Fluent there are two way of handling the turbulent boundary layer for a CFD model. One way is to suffieiently refine the mesh throughout the turbulent boundary layer so that 3 regions are eaptured in the solution. This requires that a large number of eells relative to the free stream mesh size will need to be eonstrueted. This refinement level is extremely eomputationally expensive. The alternative is more eomputationally effieient beeause it requires that a single wall adjaeent eell be eoarse enough to span the viseous sublayer and buffer regions [17]. The latter near wall meshing strategy was used for this researeh.

Either approaeh requires some knowledge about the loeal size of the turbulent boundary layer in order to appropriately separate the mesh as diseussed. To do this an equation for the non-dimensional boundary layer thickness, y^+, is used [27]:

$$y^+ = \frac{(r_w - r)u_\tau}{\nu} \tag{7.1}$$

The differenee expression in parentheses represents the wall adjaeent eell thickness, where r_w and r are the radial distanee from the axis of the nozzle to the bounds of the eell. The subscript w denotes the wall and so r_w would be the larger of the two radial dimensions. In the denominator is the local free stream kinematie viscosity and u_τ is the shear velocity, sometimes called the friction velocity, given by [22]:

$$u_\tau = \sqrt{\frac{\tau_0}{\rho}} \tag{7.2}$$

where,

$$\tau_0 = \frac{\rho u_m^2 f}{8} \tag{7.3}$$

The density, ρ, is readily available from the gas properties and the mean velocity, u_m, is approximated as the local free stream velocity. The friction factor was approximated using the aforementioned Petukhov correlation (Equation 6.6)

Equation 7.1 is rearranged so that the radial dimension of the cell can be solved for:

$$\Delta r = (r_w - r) = \frac{y^+ \nu}{u_\tau} \tag{7.4}$$

With everything else known from gas properties and/or approximated, the y^+ is the only value that needs to be determined. This is where the commitment to employing either the fine or coarse meshing strategy is important. The recommended y^+ criteria for both strategies according to Fluent is [17]:

Fine near wall meshing: $y^+ \cong 1$ (7.5)

Coarse near wall meshing: $30 \leq y^+ < 300$ (7.6)

The reasoning behind these requirements is made clear by looking at a plot of the non-dimensional turbulent velocity, u^+, versus y^+ in Figure 7.1. Recall that the fine near wall meshing strategy needs to be sufficiently refined throughout the viscous sub-layer and buffer regions. It makes sense that a lower non-dimensional boundary layer

thickness given in Equation 7.5 would correspond to a small cell height adjacent to the wall. Likewise, Equation 7.6 yields a cell height that envelopes the viscous sub-layer and the buffer regions. In addition, Fluent suggests using an upper limit for the coarse near wall meshing. In fact, for the coarse meshing strategy it is desirable to be as close to the lower limit of Equation 7.6 without going below it. If the y^+ values fall between 1 and 30 then the number of cells is said to be insufficient in the buffer region and the accuracy of the near wall treatment breaks down.

Holding the y^+ value constant at 30 in Equation 7.4, Δr becomes a function of axial position by way if its dependency on shear velocity and density. As a result there is a different wall adjacent cell height for each station. In Gambit this is captured by first creating sketch lines at key axial locations that extend the radial distance from the axis of symmetry to the wall boundary. Each line is "split" in two so that the segment close to the nozzle wall is equal to the respective near wall cell height, Δr, for that station. The other segment, which is much longer, defines the radial height where the free stream conditions are assumed to be valid. Each set of two line segments at these key locations are a kind of control line for constructing the mesh. For the base case geometry it was determined that only 3 control lines were needed to effectively split the mesh into the free stream and near wall meshes, while the aerospike required 4.

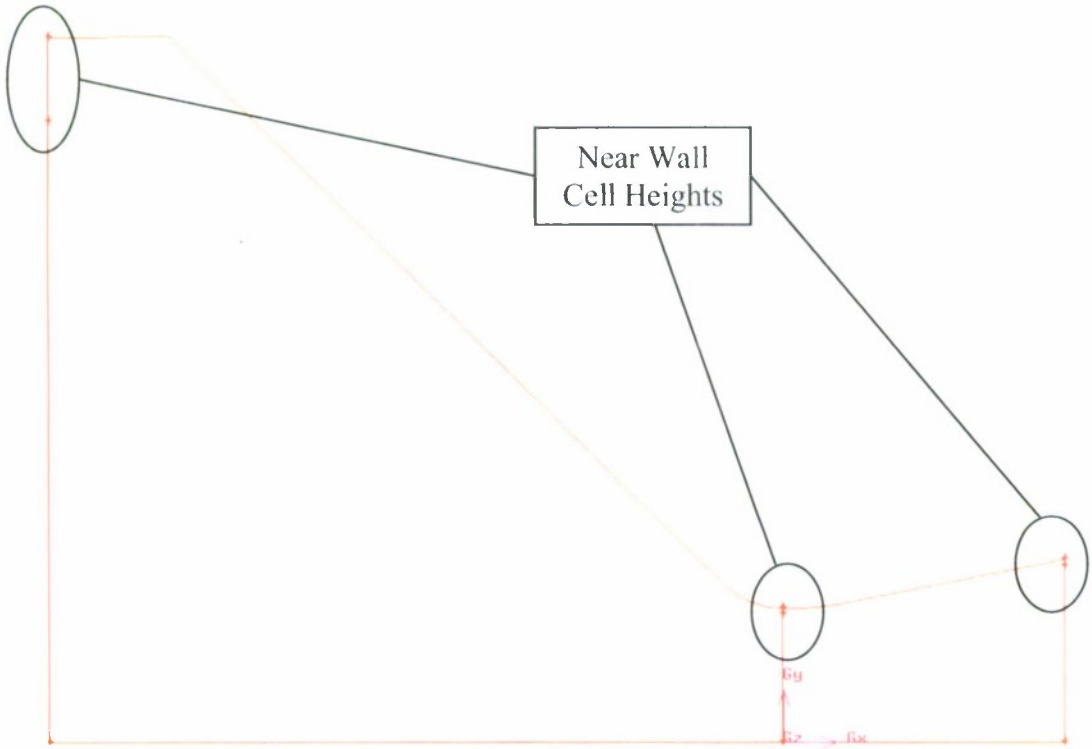


Figure 7.2 Control lines set up at the inlet, throat and exit locations for the base case study.

Since these cell heights are only approximations, they were checked by looking at the actual y^+ values that resulted when the models were executed in Fluent. This is a good way to tell if Equations 7.1-7.3 are valid. If the resulting y^+ values did not meet the criteria defined in Equations 7.5 or 7.6 then the wall adjacent cell heights were fined tuned by adjusting where the control lines are split. From here the criteria for fine tuning the mesh was based on a guess and check method. For this research, the initially calculated cell heights yielded reasonable results for the base case and only minor adjustments were made to bring the y^+ values within the scope of Equation 7.6. The aerospike geometry faired better with the initial mesh size and did not require any adjustment at all. Figure 7.3 displays an example of the structured mesh for the base case

study with the control lines applied. The separation of the free stream and wall adjacent meshes is evident near the inlet and beeomes less obvious toward the throat.

Construeting the rest of the mesh is less eomplieated. Along the axis, and in the direetion of the flow, there are 90 eells total for the base ease geometry. The same goes for the aerospike, the only differenee being that the cells are distributed along the spike wall boundary instead of the axis. 65 of these eells are in the subsonie region and 25 are in the supersonie region. Reeall that the supersonie region for the aerospike geometry is defined by a artifieial nozzle wall. When condueting the grid density study, which will be diseussed in the next ehapter, the number of eells was held eonstant in the flow direction for both eases. In the radial direetion a specified number of evenly spaeed cells in the free stream region was determined from the grid density study.

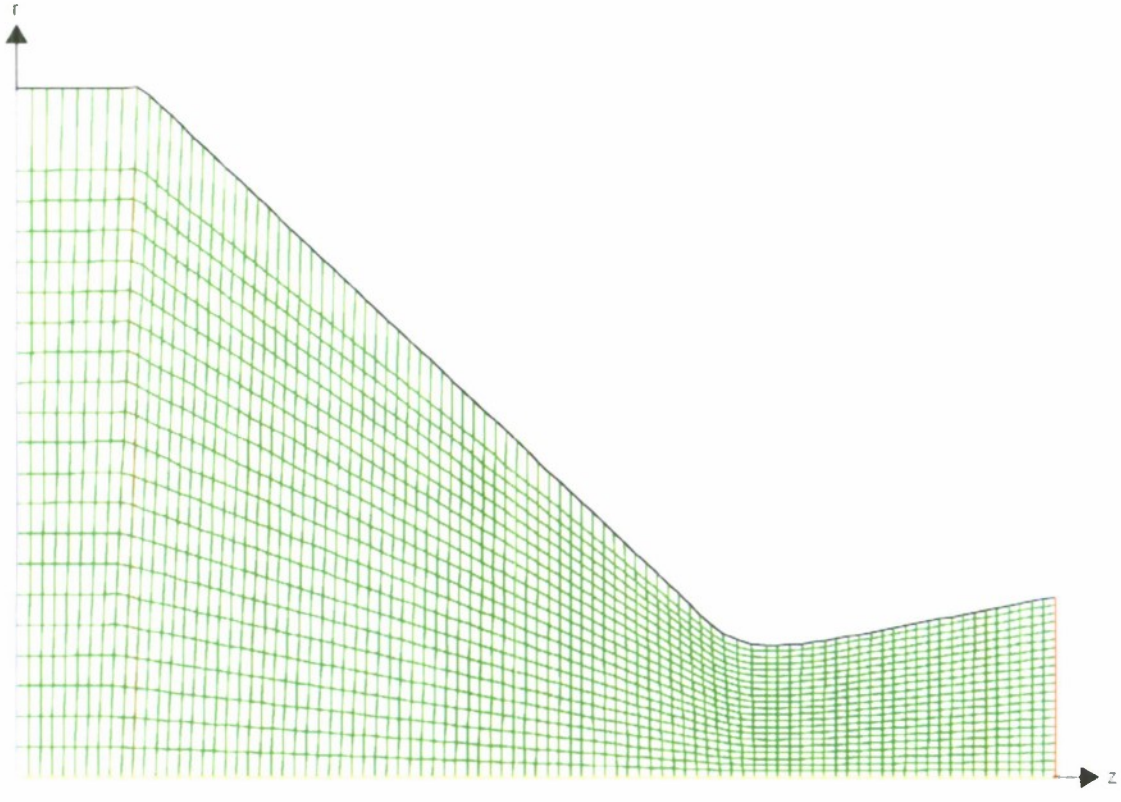


Figure 7.3 Example of the meshed base ease study geometry with 21 eells radially and 90 streamwise.

CHAPTER 8

Grid Density Study

Conducting a grid density study is important because it nearly eliminates the dependence of the numerical solution on the size of the mesh. Once this is achieved the solution more accurately reflects the effects of the boundary conditions imposed on the model. Typically, an array of densities are selected and applied to a mesh in the preprocessor. Once exported to the CFD code, a numerical model is set up and run for each mesh. The key is to maintain an identical set up for every model so that the only difference among them is the grid density. After each model is run the solutions are compared with regard to a single parameter. Ideally, the chosen parameter either increases or decreases with an increase in the grid density of the mesh. The increasing or decreasing trend will eventually converge to final value of the chosen parameter as more models are run with higher grid densities. However, it is not necessary to carry out the study to this extent. Once the slope of the trend starts to flatten out then the grid density is considered satisfactory.

For this research the base case geometry was used for the grid density study. Four meshes were used in the study with grid densities of 1890, 4590, 6390 and 8190 cells. To achieve these densities 20, 50, 70, and 90 cells were placed in the radial direction of free stream region of the mesh. The parameter used to compare these densities was an estimate for the Mach number at the exit plane of the base case geometry.

Figure 8.1 shows the results of the grid density study. The trend of the curve starts to flatten out from 6390 cells to 8190 cells. The percent difference between the Mach numbers at the exit for these grid densities is 0.024%. Based on these results the

8190 cell mesh density was used for the modeling of the base case and aerospike geometry. The final mesh density is displayed in Figure 8.2

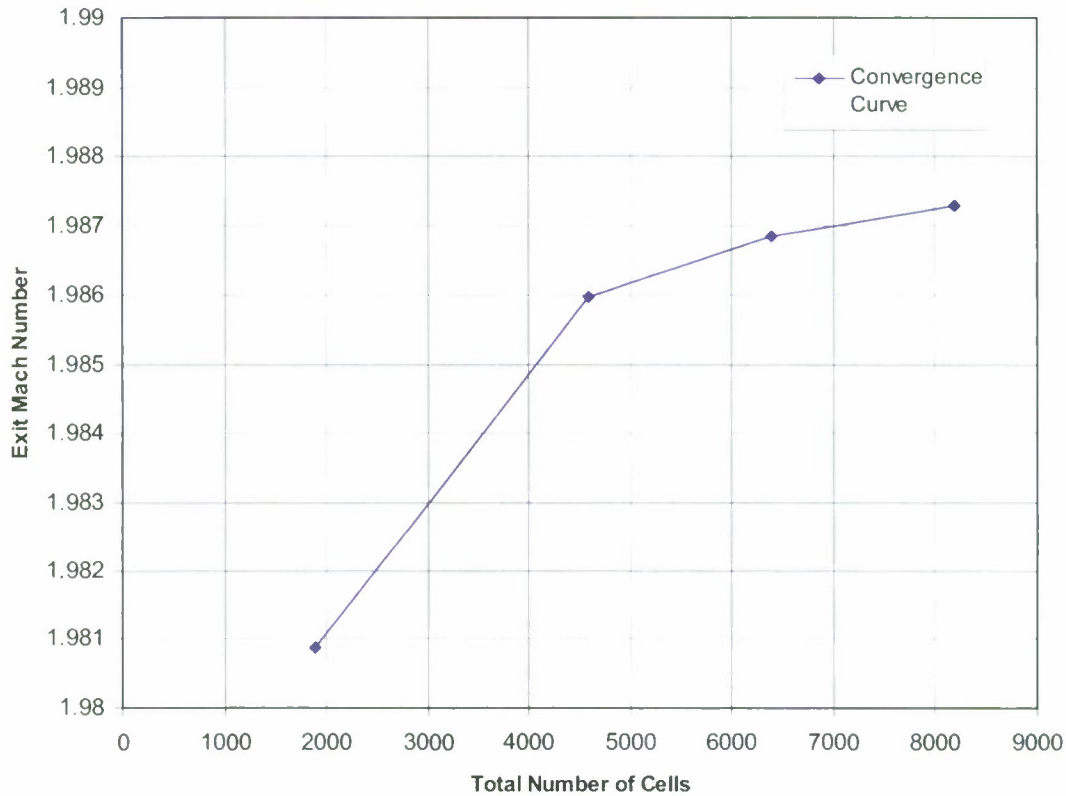


Figure 8.1 Convergence curve for the grid density study.

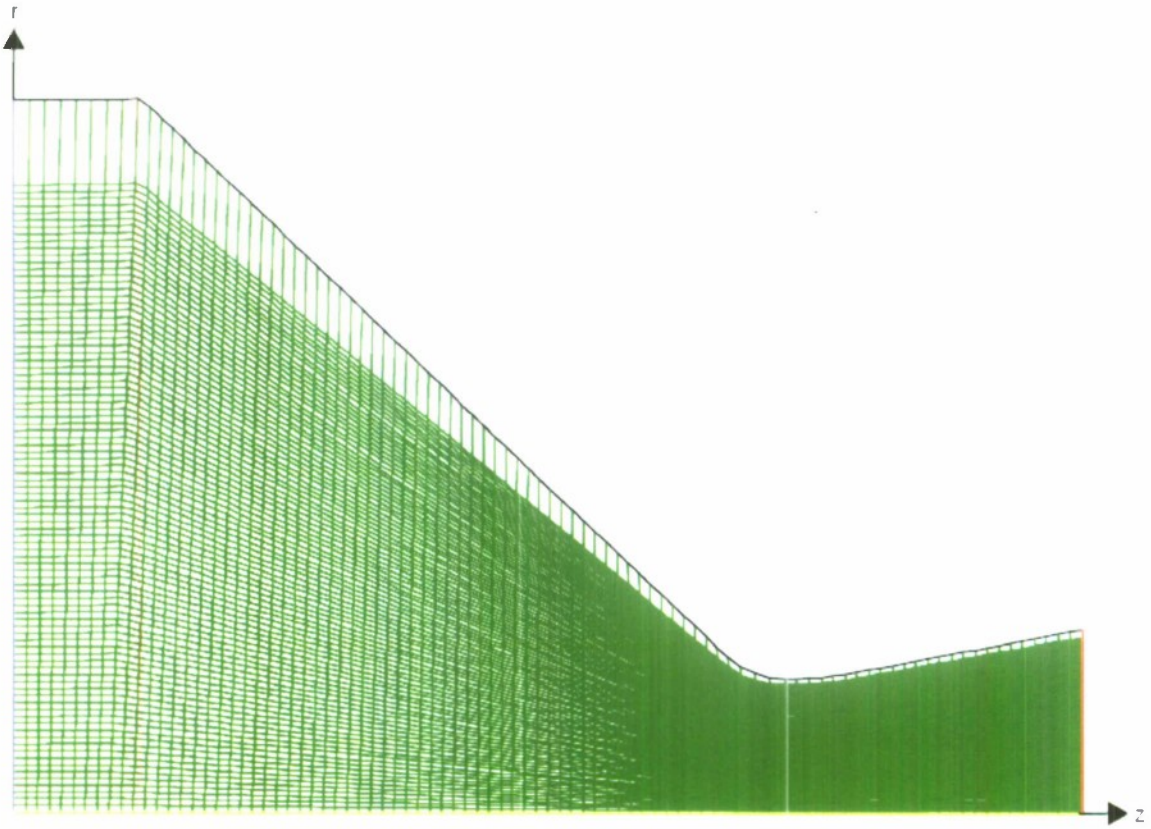


Figure 8.2 Final mesh for the base case study with 91 cells radially and 90 cells streamwise.

CHAPTER 9

Boundary Condition Seleetion

Inlet and Exit Conditions

Boundary conditions for the nozzle flow at the inlet and exit plane for each of the nozzle geometries is summarized below in Table 9.1. The inlet pressure for the base ease represents an average ehamber pressure for *"Run #20"*. For the aerospike a pressure of 300 psi was prescribed beeause it represents the intended operating chamber pressure. Both exit pressures listed eome from the eombustion analysis. Reeall that the exit plane for the aerospike nozzle has been modeled as though the nozzle hardware extends into the supersonie region when it aetually stops at the throat. The pressure listed for the exit plane is the theoretieal pressure that exists at the tip of the spike, with or without the artificial boundary in the supersonie region. The preseribed temperatures given below also eome from eombustion analysis using *TEP*.

The last two eolumns list the turbulenee parameters that were applied for the k-ε models. These initial values were ealculated using the following equations [17]:

$$\text{Turbulenee Intensity: } I = 0.16\left(\text{Re}_D\right)^{1/8} \tag{9.1}$$

$$\text{Turbulent Viseosity Ratio: } \frac{\mu_t}{\mu} = \rho C_\mu \frac{k^2}{\mu\varepsilon} \tag{9.2}$$

Where ρ is the loeal density of the fluid, k is the turbulent kinetie energy, μ is the local dynamie viseosity and ε is the turbulent dissipation rate. The turbulent kinetie energy and dissipation rate ean be approximated by [17]:

$$k = \frac{3}{2}(u_{avg} I)^2 \tag{9.3}$$

$$\varepsilon = C_{\mu}^{3/4} \frac{k^{3/2}}{l} \tag{9.4}$$

Where,

$$l = 0.07D \tag{9.5}$$

D is the local diameter of the nozzle and u_{avg} is assumed to be the local free stream velocity. The local free stream velocity is used in the absence of any good velocity profile data for the base case study. Plus, given that the goal of the base ease study is to provide a modeling approach for a new nozzle geometry for which no experimental data can exist using the free stream velocity is assumed adequate.

It is worth mentioning that Fluent suggests an expected range of 1 to 10 for the turbulent viscosity ratio. The ratio listed in table 9.1 is well above that. The significance of this large deviation from the suggested range of ratios with regard to the Standard and RNG k-ε turbulence models is discussed in susequent sections. Nonetheless, the values given, though high, are believed to be well within reason.

Table 9.1 Summary of inlet and exit boundary condition applied to CFD models in Fluent.

Nozzle Geometry	Boundary	Static Pressure, MPa *(psi)*	Static Temperature, K *(°R)*	Turbulence Intensity	Turbulent Viscosity Ratio
Base Case	Inlet	1.498 *(217.20)*	2986 *(5374.8)*	4.28%	164
	Exit	.208 *(30.17)*	2310 *(4158)*	3.55%	283
Aerospike	Inlet	2.068 *(300.00)*	3178 *(5721)*	4.22%	170
	Exit	.103 *(14.99)*	2086 *(3754.8)*	3.51%	619

Nozzle Wall

The fully developed turbulent flow assumption that justifies assigning a constant temperature along the nozzle wall was tested to determine its validity. During the base

ease study two different temperatures were assigned to the nozzle wall. The first temperature choice was made by simply leaving the default value in Fluent at 300K (*540°R*) when imposing the constant wall temperature condition. The second value was chosen somewhat arbitrarily as 1900K (*3420°R*), although there is some logic behind it. It is expected that 1900K is on the same order of magnitude as the actual wall temperatures for both nozzles. Yet, it is lower than the lowest static temperature predicted in the combustion analysis eliminating confusion of a potential for heat fluxes predicted by Fluent to change sign. Also, the chosen temperatures differ by 1600K which should make obvious any differences in resulting Nusselt numbers.

CHAPTER 10

The Aerospike Geometry

Most of the work in this portion of the researeh has been dedicated to conducting an effeetive base ease study. The fruits of this effort were the relative ease in applying the same strategies to the aerospike geometry. Referring back to Figure 6.3, the aerospike study was mapped out in a similar fashion but without the need for semi-empirieal eorrelations or determining the best turbulenee model. There are, however, some signifieant differenees direetly related to the aerospike geometry worth mentioning with regard to how these strategies were implemented.

Flow through the aerospike nozzle has an upper and lower bound defined by hardware, whereas the conventional nozzle only has hardware at the upper bound. This differenee made the station lines more difficult to set up for the aerospike. This is beeause station lines ean longer be eonstrueted normal to the axis of revolution like with the conventional nozzle. Instead, each station line was positioned approximately normal to the spike eontour. In some sense, this approaeh follows the same form as the stations set up for the eonventional nozzle beeause lines are still plaeed normal to the lower bound, it just happens the lower bound of the aerospike is a two dimensional eontour and not a straight line. It was the intent of this researeh to eonstruet station lines in this way to provide the best level of eonsisteney between both nozzle geometries with regard to monitoring the flow.

In doing so, the eonvenienee of evenly spaeing eaeh station was lost, as well as the parallelism. The same number of stations was used throughout the subsonie region as

in the base case study. However, the overall number of stations was reduced because the artificial supersonic region was neglected.

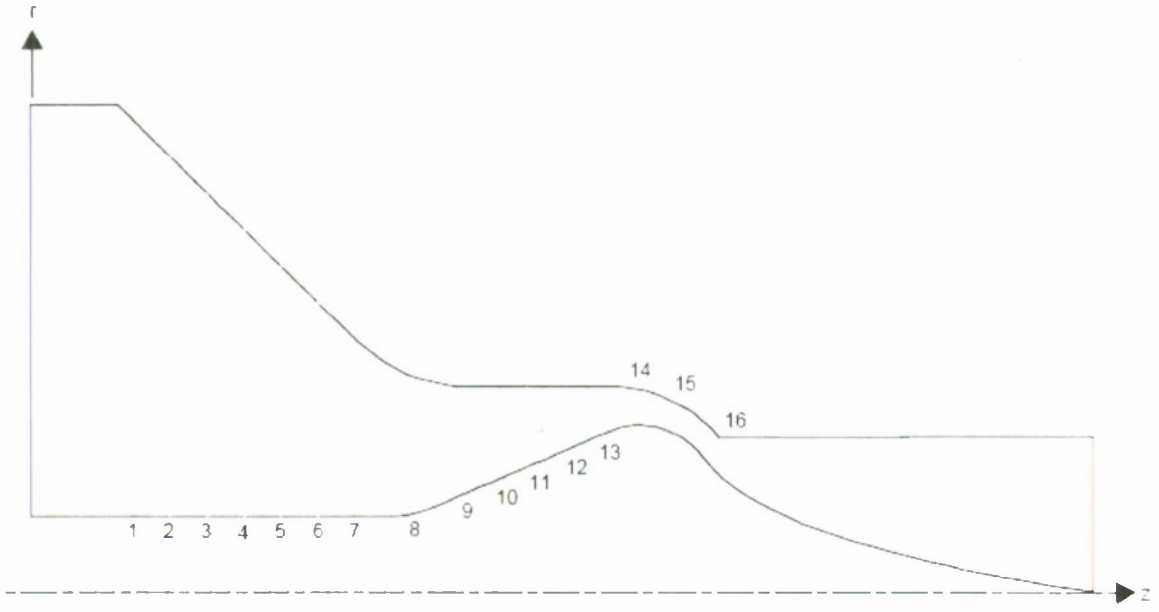


Figure 10.1 Stations defined for the aerospike geometry.

Another convenience that is lost for the aerospike is the geometric simplicity of the cross-sectional areas taken at each station. For stations lines 1 through 7 a position normal to the spike results in a 90° orientation to the axis of symmetry. The calculation for the area of annulus is a simple difference in the circular areas bounded by the local outer and inner radii:

$$A_{station} = \pi(r_o^{\,2} - r_i^{\,2}) \tag{10.1}$$

Where r_o and r_i represent the radial locations of the outer and inner station end points, respectively. Similarly, the hydraulic diameter reduces to a difference between the outer and inner diameters:

$$D_H = \frac{2A_{station}}{\pi(r_o + r_i)} = 2(r_o - r_i) \tag{10.2}$$

For stations 8 through 13 a normal position relative to the spike surface creates an obtuse orientation with the axis of rotation and for stations 14 through 16 one that is acute. This makes calculating hydraulic diameters more difficult because the cross sectional area is no longer planar. The area of the annulus is defined as a surface area for a conical section and is given by:

$$A_{station} = \pi\left[\left(r_o - r_i\right)\left(\sqrt{\left(\frac{r_i}{\tan\phi}\right)^2 + r_i^2}\right) + r_o\sqrt{\left(r_o - r_i\right)^2 + \left(z_o - z_i\right)^2}\right] \tag{10.3}$$

Where z_o and z_i represent the axial locations of the outer and inner station end points, respectively. The angle ϕ is taken as the smallest supplementary angle made between a station line and the axis of rotation. To approximate an effective hydraulic diameter the area of the station is assumed to be a planar annulus with an inner radius equal to r_i. In this way Equation 10.2 can be applied by first solving for an effective outer radius for the assumed planar annulus. The final formulation is:

$$D_H = \frac{2A_{station}}{\pi\left(\sqrt{\frac{A_{station} + \pi r_i^2}{\pi}} + r_i\right)} \tag{10.4}$$

The use of these hydraulic diameters is not critical to analyzing results for the CFD model. Recall that these coefficients are determined directly using Newton's Law of Cooling. However, the inclusion of these hydraulic diameters supports an attempt to compare the CFD results for the aerospike geometry to the semi-empirical correlations and to do so requires that the coefficients be non-dimensionalized into local Nusselt numbers. Equations 10.1 through 10.4 were used to provide hydraulic diameters for that

non-dimensionalization. Gas properties specific for the aerospike geometry and the intended operating conditions were approximated using *TEP*, on the basis of area ratio. A list of these properties can be found in Appendix C.

The 4 control lines used for the wall adjacent cell definition were placed at the inlet and stations 7, 12 and 16(Figure 10.1). Stations 7 and 12 are important for mapping the wall adjacent cells around the curvature of the nozzle wall. Station 16 is at the throat and it only makes sense to control the cell spacing here since this is were the highest convective heat transfer coefficient is expected.

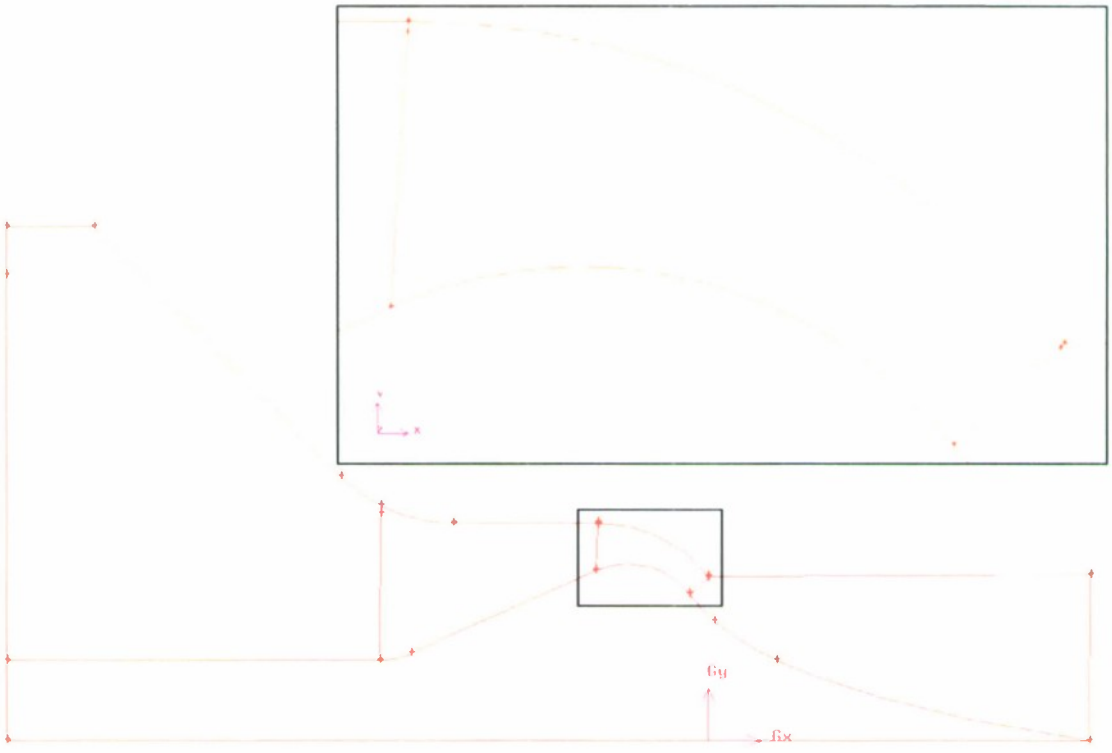


Figure 10.2 Control lines at the inlet, throat and two positions in between for the aerospike geometry.

There are some potential pitfalls with this handling of the geometry and station set up. When data is retrieved from Fluent for analysis it is difficult to associate outputs with axial position for those stations that are positioned at an angle. For instance, in

order to calculate local Nusselt numbers a local mean temperature of the fluid is required. For those stations that are situated at an angle to the axial coordinate the temperature outputs will represent an average value over a specified range of axial locations as well as radial locations. So, it becomes difficult to discuss "local" properties of the fluid relative to one axial coordinate. For this research the axial location of the midpoint of each station was adopted as the "local" coordinate.

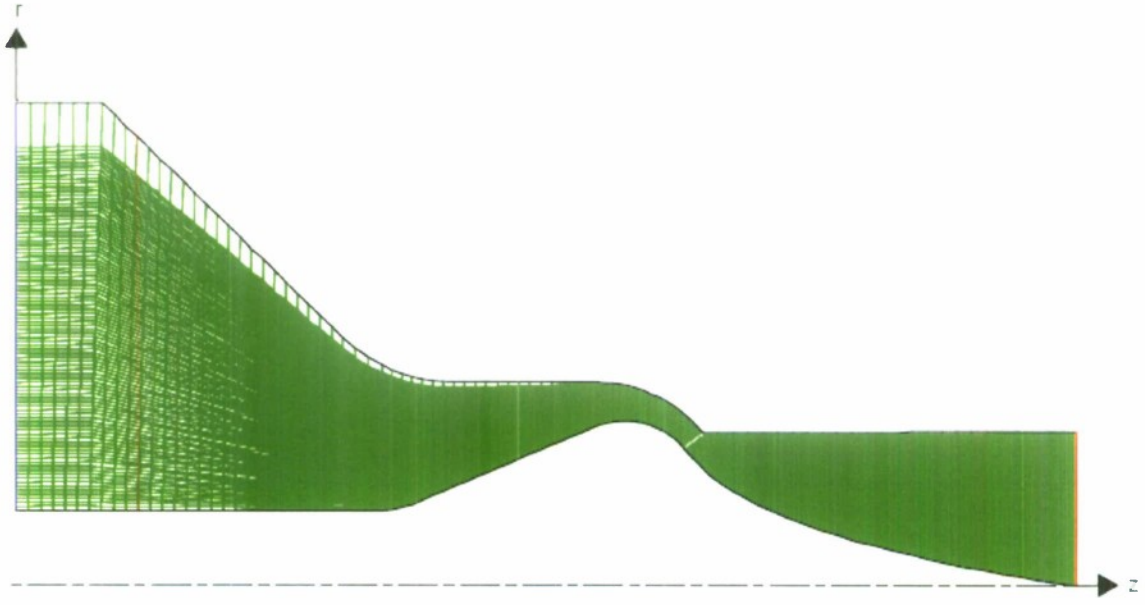


Figure 10.3 Final mesh for the aerospike geometry having 91 cells radially and 90 streamwise.

CHAPTER 11

CFD Model Results and Discussion

The results of this CFD study first explore the base case study with regard to the *Standard* and *RNG k- ε* turbulence models. Results for two constant wall temperatures follow that display the reasonability of the fully developed turbulent flow assumption. Finally, results for the aerospike geometry are given and are supported by the verification of the base case study. All results are compared with corresponding results using the four semi-empirical correlations. In addition, outlines of the nozzle wall are provided for both geometries to give some idea of where the results are located in the hardware.

Figure 11.1 shows both models overlaid on top of the semi-empirical correlations. An outline of the nozzle wall for the base case study has been provided to give an idea of where the throat of the nozzle exists. As expected, the highest Nusselt number for both models is near the throat. The most obvious distinction between these models is the magnitude of the Nusselt number at the throat for the *Standard k- ε* model. This model significantly over-predicts the heat transfer at the throat of the nozzle and for much of rest of the nozzle, too. Near the throat the *Standard k-ε* model yields a Nusselt number that is 79% higher than the average of the correlations at the same point. At its highest Nusselt number prediction the *Standard k-ε* model deviates from Long's formulation, the more conservative of the four correlations, by approximately 70%.

The *RNG k- ε* model displays better agreement against the correlations, following especially close to the Bartz and Greenfield formulations for much of the subsonic region. At the throat the *RNG k-ε* model deviates from the average correlation Nusselt number by only 6%. The largest deviation in Nusselt number was found to be 34% with

respect to Long's formulation just before the throat in the subsonic region. Throughout the subsonic region there is a general trend of a slight under-prediction and the converse seems true for the supersonic region. In addition, the shape of the curve in the supersonic region is somewhat erratic. This poses little concern since the aerospike model will neglect the heat transfer effects in the supersonic region for lack of actual hardware.

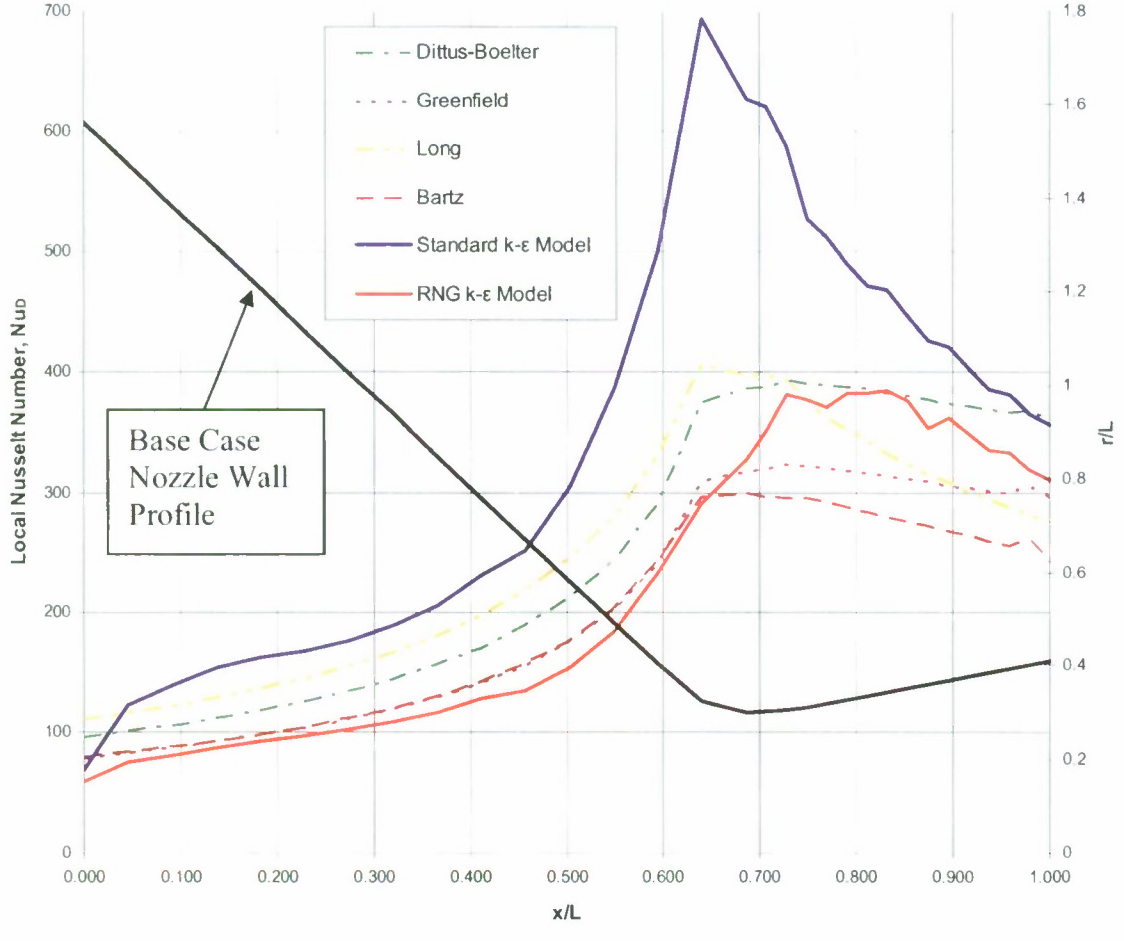


Figure 11.1 Local Nusselt number plot comparing Standard and RNG k-ε models to semi-empirical correlations.

The results of the sensitivity study of the Nusselt numbers predictions with a change in the prescribed constant wall temperature are displayed in Figure 11.2. Nusselt number curves for both wall temperatures have the exact same trends and shape, even with respect to the non-smooth trend exhibited in the supersonic region, but with slightly

different magnitudes. In theory, any disagreement between Nusselt numbers for these two constant wall temperatures suggests that the flow is not fully developed turbulent flow. However, that does not mean that such an assumption is unreasonable. These results deviate from each other about as much as they deviate from the correlations. At their largest deviation, the T_w = 300K curve predicts Nusselt number values that are 11% higher than the T_w = 1900 K curve. Consider that a 300 K specification at the wall would yield a large thermal gradient through the thermal boundary layer. The good agreement exhibited between the 300K wall specification and the 1900 K wall specification would suggest that even though the flow is not theoretically fully developed it must be really close. Qualitatively, the results support the assumptions of fully developed turbulent flow for practical use in predicting Nusselt numbers in the design of the aerospike nozzle. The 1900 K wall temperature was used for the aerospike CFD model because it is believed to more closely represent the actual wall temperatures than a 300 K specification.

Figure 11.3 shows the final Nusselt number results for the aerospike CFD model. As a check the correlations were included for comparison based on the gas properties for the intended operating conditions of this nozzle configuration. Recall that one of the pitfalls in setting up the stations for the aerospike model was that displaying results with regard to the axial coordinate would be cumbersome with regard to the orientation of the station lines. This is evident in Figure 11.3. Unlike the base case study, the Nusselt number distribution peaks before the throat, between x/L= 0.4 and 0.5. The trend decreases through x/L= 0.5 and 0.6 before peaking again at the throat. It would seem that the CFD model is not adequately predicting the correct Nusselt numbers. However, the

correlations show the same trends. Since the thermal conductivity of the gas is held constant the explanation has to lie with the hydraulic diameters used for each station or with the convective heat transfer coefficients.

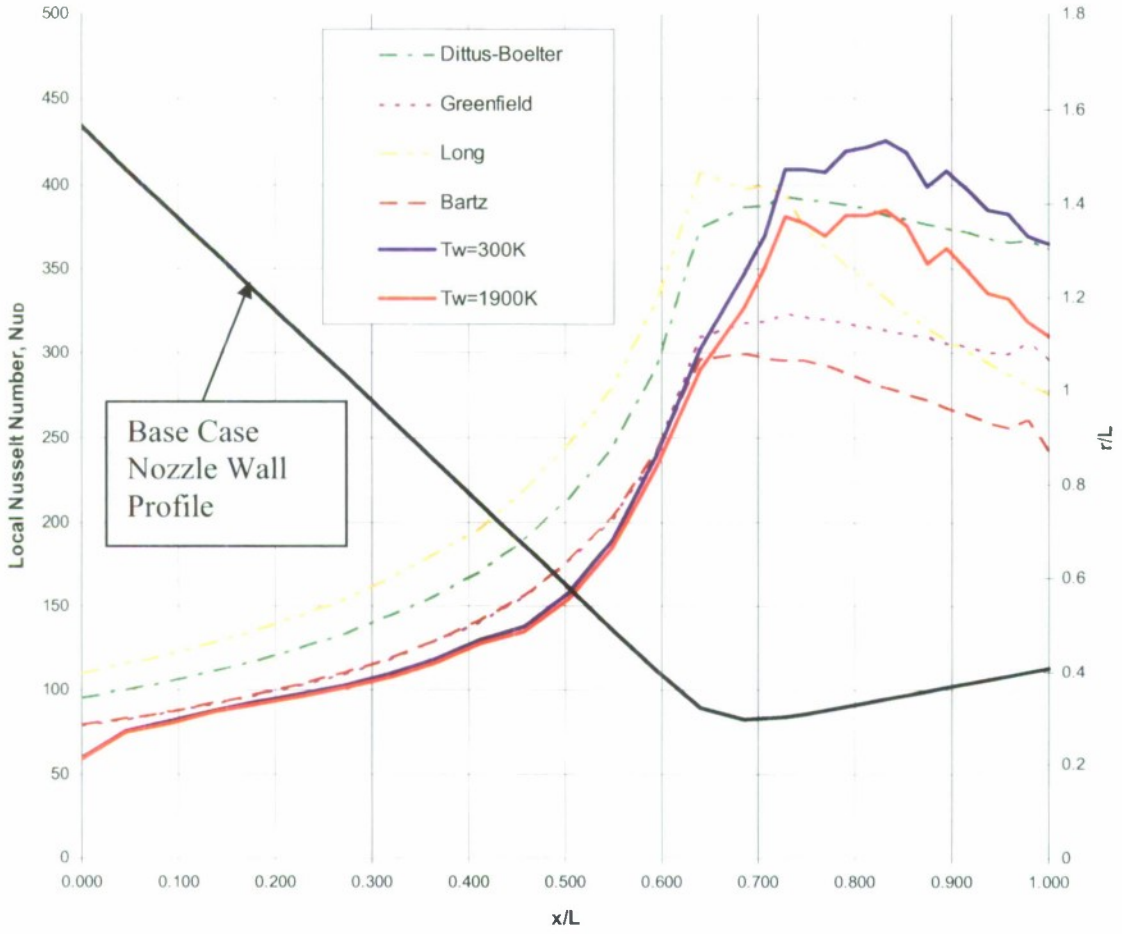


Figure 11.2 Local Nusselt number plot comparing results for RNG k-ε models using different constant wall temperatures to the semi-empirical correlations.

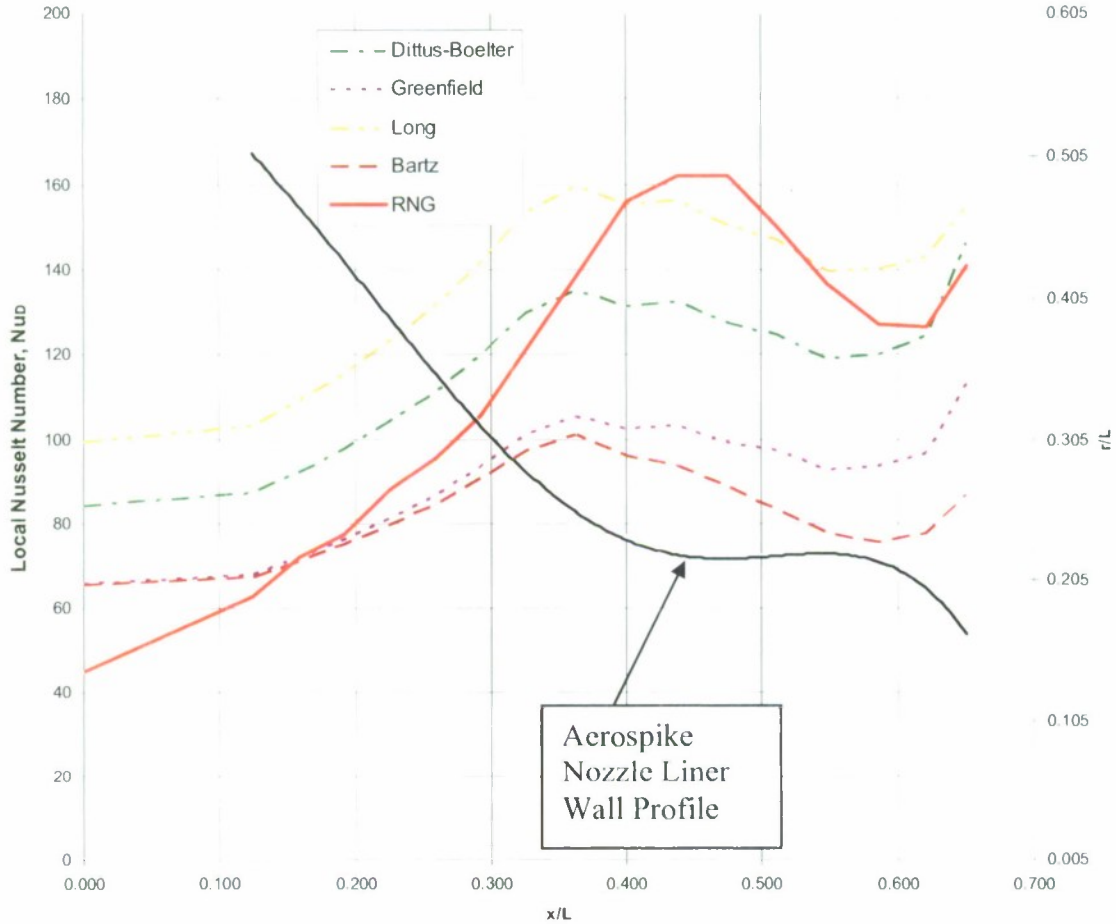


Figure 11.3 Local Nusselt number plot using RNG k-ε model for the aerospike geometry. Semi-empirical correlations are displayed for perspective.

A plot of the convective heat transfer coefficient against axial position is given in Figure 11.4. The trend shown here is more intuitive with regards to the traditional expectations of the highest heat transfer rate existing at the throat of a rocket nozzle. Furthermore, there is very reasonable agreement between coefficients predicted by the *RNG* model and the semi-empirical correlations typically used for conventional nozzles. What this suggests is that the discrepancy between Figures 11.3 and 11.4 is most likely related to the hydraulic diameters used to define the local Nusselt number. Recall that all of the station lines were constructed normal to the surface of the spike which forced some of them to be at angles relative to the axis of symmetry. As a result some of the hydraulic

diameters were ealeulated using a modified form of the Equation 10.2. This modification will need to be reviewed or replaced by a more appropriate method in future studies. However, the integrity of the CFD is not eompromised by this modifieation. The results of the CFD model have no dependeney on these effeetive hydraulie diameters. The intent underlying the use of hydraulic diameters was to provide a non-dimensionalized eomparison to the correlations in an effort to display a possible usefulness of these correlations in lieu of a CFD model. In the end the direct comparison of convective coeffieients satisfies this effort.

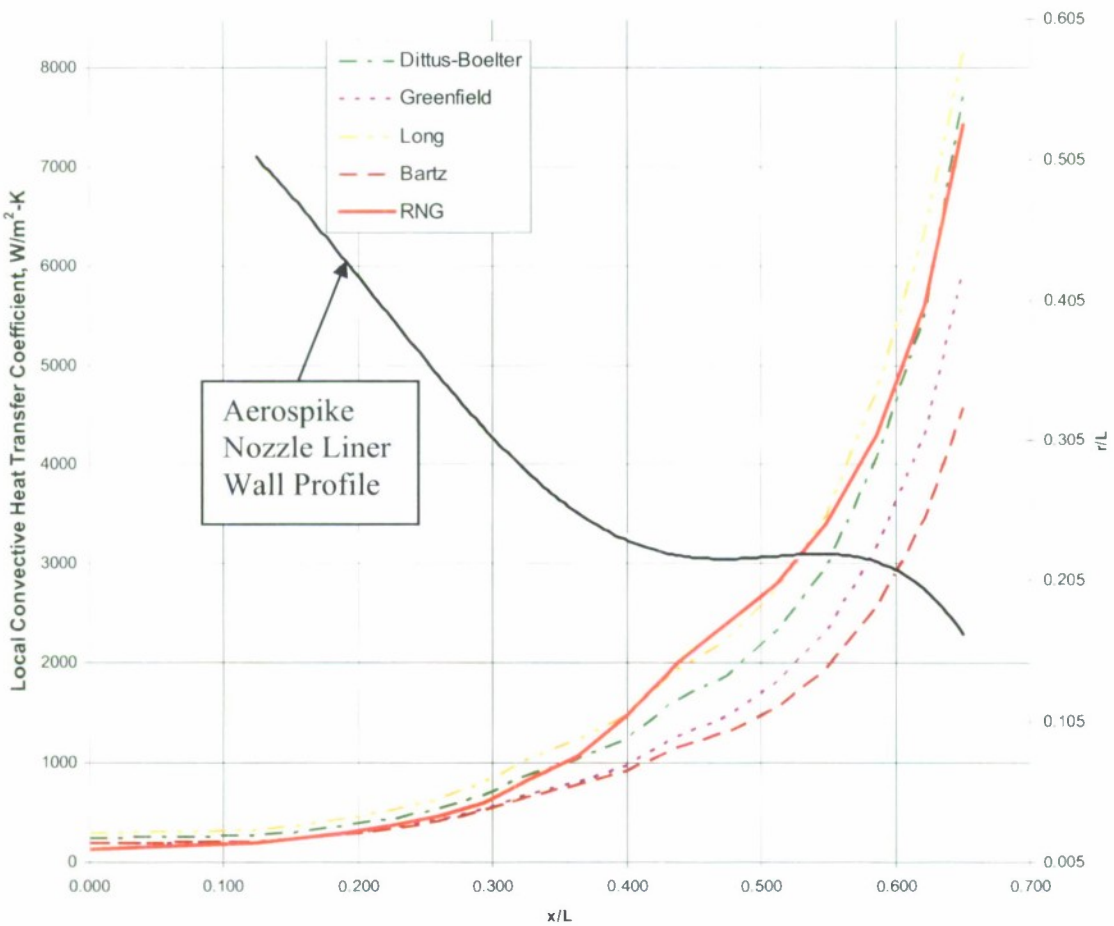


Figure 11.4 Loeal eeonveetive heat transfer eoefficient plot versus axial loeation using RNG k-ε model for the aerospike geometry. Semi-empirical eorrelations are displayed for perspeetive.

The CFD modeling efforts have produced a strategy for predicting steady state gas-side heat transfer coefficients for the aerospike nozzle. The base case study served as a verification tool for the final aerospike study. In the base case study assumptions were made about the meshing scheme near the wall and the reasonability of the fully developed turbulent flow. The fully developed flow assumption has been deemed reasonable for practical use in the aerospike design effort. The meshing scheme used to capture the heat transfer effects has been verified as well. Given the large uncertainties associated with turbulent flow, the use of the *RNG* k-ε model is considered a successful numerical tool for predicting local convective heat transfer coefficients for the aerospike geometry. The next phase of this design study relied on the results of this CFD model to predict gas-side wall temperatures and to develop the thermal structural model.

CHAPTER 12

FEA Modeling Approach

The results of the CFD study were used in conjunction with various materials to complete the design study of an un-cooled annular aerospike nozzle using Abaqus Standard. The modeling approach in Abaqus has been split up into two distinct parts: A transient heat conduction model and a thermal structural model. The latter is dependent on the nodal temperatures produced in the former. The existence of the spokes in the spike web made it difficult to model it along with the rest of the assembly and so a separate study was conducted. In addition, this spike web was designed for an assumed lateral loading case independent of the thermal structural model. The spike web will be discussed in later chapters, but suffice to say that modeling strategies developed here will translate.

Unit Specification

Abaqus is designed to be universally applicable for any system of units and so there is no option to distinguish between Standard and SI system. Instead, a system of units is prescribed through user inputs for quantities like material properties, loads, boundary conditions. For the assembly and spike web modeling efforts these inputs were specified in terms of meters(m), Kelvin(K), kilogram(kg), Joules(J), Pascals(Pa) and seconds(s).

Meshing

Identical model geometries and grid densities were used for both the transient heat conduction model and the thermal structural model to aide the importation of resulting temperatures into the structural model. The mesh was set up for all parts of the

FEA study using quadratic, quadrilateral cells. For the transient heat conduction models heat transfer elements were specified. Axisymmetric stress elements were specified for the thermal structural models and the final spike web lateral loading model. A brief grid density study was conducted for both the heat conduction and thermal structural. A global seeding size of 0.00125 m was determined to be sufficient for the 2-D axisymmetric assembly. For spike web the global seeding size was also set to 0.00125 m, except for the lateral loading case where the seeding size was 0.000625 m.

Verification Study

A verification study was conducted to ensure that numerical methods in Abaqus yielded results for simple problems that could be replicated using analytical methods. This is an important step when using any FEA code and is similar in philosophy to the base case study conducted for the CFD model. By conducting this study a greater level of confidence about the features utilized in Abaqus is established. In addition, these studies can help identify human errors related to material inputs, user subroutines, boundary conditions, etc. The verification study was split up to focus on four analysis features used for the transient heat conduction and thermal structural models.

Two simple axisymmetric conduction heat transfer problems were used to verify steady state and transient heat transfer capabilities in Abaqus Standard. A simple cylinder was used for the steady state conduction problem with an inner and outer diameter within an order of magnitude of the dimensions for the aerospike nozzle. The transient heat conduction was set up using a solid cylinder, also scaled within the same order of magnitude of the nozzle. The inclusion of the steady state conduction may seem redundant considering that a transient conduction problem is also studied. In truth, the

purpose of steady state eonduetion study was used to better understand how to effectively map temperatures into a thermal struetural analysis, whieh represents the third problem in the verifieation study.

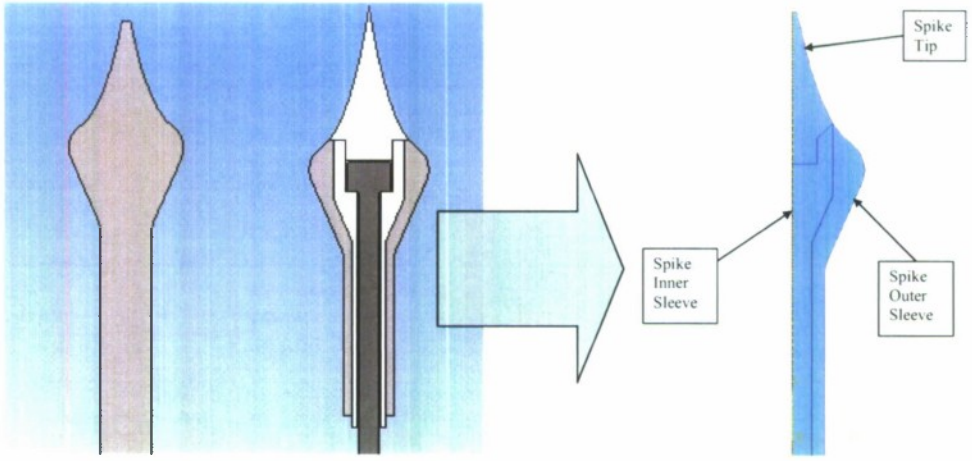
The third problem dealt with analyzing thermal stresses for a simple hollow eylinder. The cylindrieal model used was made identical to the one used for the steady state conduction problem so that temperatures could be mapped in to the same nodes. Thermal stresses that result from the mapped in temperatures were eompared to analytieal results. The fourth problem for the verification study uses the same simple eylinder geometry, but with an applied internal and external pressure in lieu of a temperature gradient. This model was also verified using analytieal results. A series of sample ealeulations that support this verifieation study ean be found in Appendix F.

Modeling the Geometry

The geometry of the aerospike nozzle was modeled as a 2-D axisymmetric assembly of eomponents in Abaqus using a similar modeling approaeh to the CFD study. The only differenee is that in the CFD study there was only one part geometry that represented the flow field between the spike and the nozzle liner. For this study seven part geometries were imported from Solidworks: the spike, nozzle liner, nozzle outer, spike web (outer rim only), gasket, ehamber outer and ehamber inner. The inelusion of parts like the ehamber outer, ehamber liner and gasket that exist upstream of the nozzle on the apparatus were ineluded so that the effects of their interaetion with the nozzle eomponents eould be eaptured. However, the presenee of the inner and outer fuel grains was exeluded, along with any of the insulating benefits that they may provide. The idea behind their exelusion is that modeling exposed ehamber hardware represents a worst

case scenario. This turned out to be of little consequence as the film coefficients in the chamber are comparatively low and yield relatively low heat transfer rates in the chamber hardware.

Part of this study will investigate two proposed designs for the spike. Several different material configurations for the construction of the spike were modeled in Abaqus based on two concepts that are shown in Figure 12.1. One concept is made up entirely of one material (Figure 12.1a). This exhibits the benefit of a simple construction but creates complications for assembly and serviceability. The second concept boasts a logical approach to being integrated into the rest of the assembly but makes for a more complex construction (Figure 12.1b). This approach separates the spike into two pieces; a main body and a spike tip. The main body is a semi-permanent construction of an outer and inner sleeve. The idea is that the outer sleeve would see the harshest conditions while the inner sleeve would take the axial load of the spike and provide mounting provisions for the spike tip. A threaded fastener is located at the center of the inner sleeve that is accessible by removing the spike tip. This fastener makes repeated disassembly possible which is a desirable attribute for labscale testing.



(a) (b) (c)

Figure 12.1 Renderings of both spike concepts and the as-modeled geometry.

Both concepts are represented by one model in Abaqus and are distinguished by assigning different material properties to the appropriate regions shown in Figure 12.1c. The existence of the fastener is absorbed into the inner sleeve, treating it as a solid core. The spike tip has been altered slightly to make the mating surfaces between the tip and inner sleeve normal to the flow of gases.

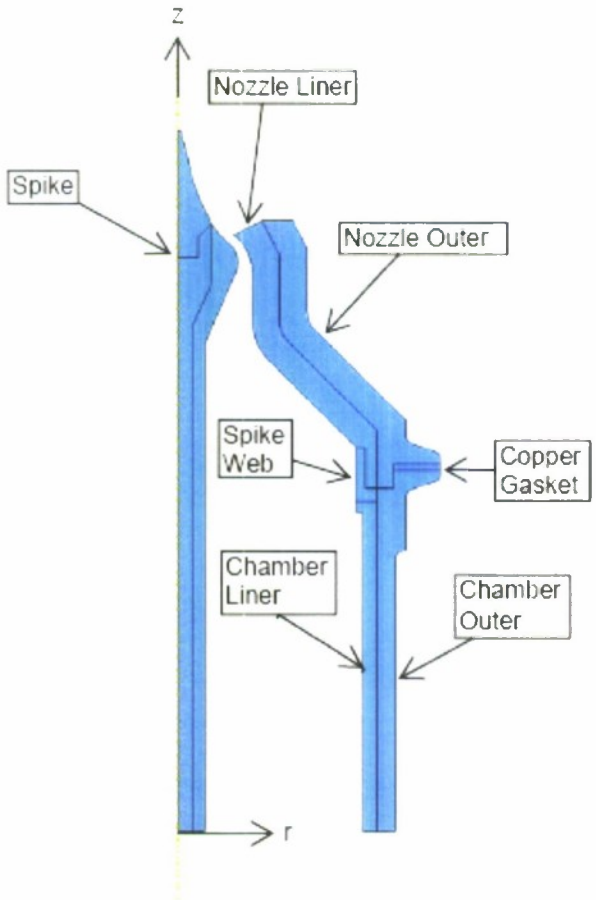


Figure 12.2 The spike assembly as modeled in Abaqus.

Material Properties

Material properties for the selections outlined in Chapter 2 were input into Abaqus Standard. In most cases basic material property data was available to effectively characterize each selection mechanically and thermally. Where necessary, assumptions were made to complete a set of properties for a material if information was limited. In addition, an effort was made to use temperature-dependent properties when available to increase the accuracy of the heat conduction and thermal structural models. A complete set of material properties can be found in Appendix M.

For the selected alloys this temperature-dependent data was pulled directly from a specialized design allowables handbook [8]. Isotropic material constants were approximated for a range of temperatures from available curves in the handbook. Physical properties for specific heat, thermal conductivity, coefficient of thermal expansion and density were approximated from similar curves. Yield and ultimate stress data was also pulled from this handbook and used in the analysis as a function of temperature. Similar data was pulled from this handbook for the copper gasket.

For Columbium FS85 and 304L stainless steel it became necessary to consider plastic behavior. To support this non-linear behavior plastic hardening data was input for these materials in addition to elastic constants. Plastic hardening data was input for the stainless steel at two temperatures [28]. A linear slope was approximated for each hardening curve and used to assemble plastic hardening estimates for the Columbium. To make use of these non-linear mechanical properties the non-linear option was turned on for relevant steps during the analysis and is emphasized in later chapters. All other materials used in this research were assumed to exhibit linear-elastic behavior.

Unfortunately, temperature-dependent information was not available for the grade SiC-6 graphite and so one set of data from a material supplier was used at all temperatures [9]. To model the C/SiC composite preform material data came from various sources because information is limited for the specific material that is to be supplied by *Hypertherm*. Published work was interpreted to produce a set of temperature-dependant orthotropic material constants. Other physical properties were estimated using available laminate data[16].

Theoretical estimates for the stiffness matrix produced by Dalmaz *et al.* for 2.5D C/SiC were interpreted to provide properties at two temperatures [15]. For C/SiC the temperature at which the stiffnesses are highest is the same temperature at which the silicon carbide matrix is infiltrated into the carbon fiber weaves. When the matrix is applied at this temperature, the composite is in an expanded state and so there are no residual stresses between it and the fibers. However, as the composite cools to room temperature residual stresses form and tiny cracks are created that reduce the stiffness properties. The work of Dalmaz *et al.* offers one stiffness matrix that exhibits similarities to empirical data taken at room temperature and another that is considered "ideally non-cracked". These two stiffness matrices were applied to represent room temperature (300 K) and high temperature (2273 K) properties for the 2.5D C/SiC available through *Hypertherm*. Relevant temperature-dependent data was not available for other properties and were assumed to be constant for all temperatures.

Since the aforementioned material stiffness properties are orthotropic they were input into Abaqus as such. However, the way in which they were assigned to the model is not necessarily representative of actual hardware and requires some explanation. An

actual C/SiC composite spike part produced by *Hypertherm* would start out as a preform. The preform is essentially a rectangular shaped brick and has fibers in three orthogonal directions. This aligns the orthotropic properties with a Cartesian coordinate system. The preform would then be machined into the shape of the spike using specialized diamond coated tooling. In the end, the finished part would have an axisymmetric geometry but orthogonal fiber directions. This means that in reality the spike will likely not exhibit axisymmetric behavior when loaded thermally or otherwise. This presents a problem in modeling the spike as a 2-D axisymmetric part using these properties. When the properties are applied to the two-dimensional planar part in Figure 12.2 they will be orthotropic with regard to a Cartesian coordinate system in that one plane. But, since the geometry is axisymmetric, those properties are "wrapped" around the axis of symmetry making them orthotropic with regard to a cylindrical coordinate system. Figure 12.4 offers a visual representation of the as-delivered and as-modeled fiber directions for the 2.5D C/SiC composite. The impact that this misrepresentation of orthotropic material properties has on the results is discussed later.

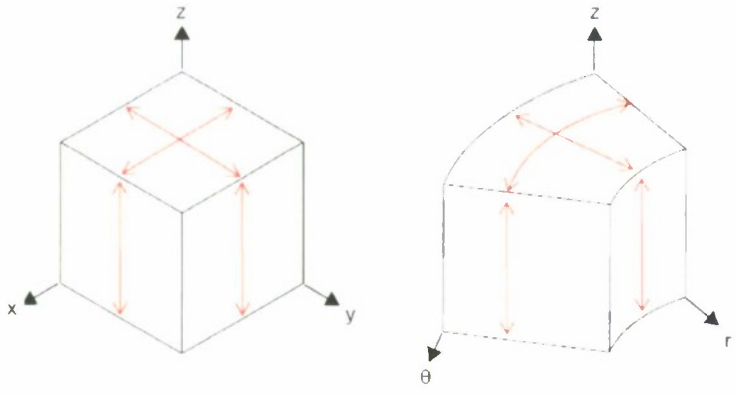


Figure 12.3 Fiber directions for the 2.5D C/SiC material.

CHAPTER 13

Transient Heat Conduction Model

Using the material properties and geometry, a transient heat conduction model was set up to simulate how heat is transferred from the combusted gases to the components of the assembly. Boundary conditions for this model were specified in Abaqus using two types of contact interactions. One defines the gap conductance between any two mating surfaces and the other applies a film coefficient to surfaces exposed to the hot gases and/or ambient air. These interactions are implemented for periods of time defined through steps in Abaqus. Four steps were used to approximately simulate a test run of a hot-fire scenario, although the first two steps define conditions at $t=0$. The first step is called the "initial step" and was used to assign all of the surface-to-surface gap conductances. The ambient conditions are defined in the next step (Step 1). In the next two steps (Steps 2 and 3) the film coefficients developed in the CFD model are applied to the interior surfaces and are made dependent on time.

At the beginning (Step 1) of the simulated run, the exposed inner and outer surfaces of the assembly are subjected to ambient conditions. The inner surface is defined by the exposed surface of the spike, nozzle liner, spike web and chamber liner. The outer surface is defined by the exposed surface of the nozzle outer and chamber outer. The ambient conditions are characterized using a film coefficient of 5 W/m^2-K and an approximate room temperature of 300K. The specified film coefficient represents a generally accepted value for approximating natural convection [22]. It should be noted that effects of radiation heat transfer from the outer surface to the ambient environment have neglected, but are significant. Steps 2 and 3 change the definition of the film

eoeffieient applied at the inner surfaees, but propagates the conditions for other surfaces from previous steps. The new film eoefficient definition is characterized for the hot gases that flow through the chamber and nozzle.

Convective coeffieients from the CFD study developed for the surface of the nozzle liner were compiled into an Abaqus FILM user-subroutine along with respeetive axial locations and adiabatic wall temperatures [29]. The user-subroutine is set up to distribute interpolated film coefficients and sink (adiabatic wall) temperatures to the inner surfaces based on the axial eoordinate of the nodes that lie on that surfaee. Code for this user-subroutine ean be found in Appendix E. It should be noted that the resulting conveetive heat transfer coefficients from the CFD study assume steady state flow but are here applied to a transient heat eonduetion model. In addition, these film eoefficients are being applied to the spike and chamber liner surfaces as well as the nozzle liner. This requires that certain assumptions be made for their application to the transient heat transfer model [30].

In the CFD model the spike surface was defined as adiabatie to meet the definition of an un-cooled design, and because the model assumed steady state eonditions. As a result no eonveetive heat transfer eoeffieients were developed for the spike surface. Also, film coeffieients were not developed for the inner surfaces of the chamber hardware since the CFD model did not eonsider these regions. However, there is a desire to simulate conveetive heat transfer at these surfaees for a transient eonduetion model to understand how the model behaves when different materials are applied. To meet this need the heat transfer eoeffieients developed for the nozzle liner were to other surfaces through a series of assumptions. Figure 13.1 shows the inner surfaees of the

assembly model split up into 5 regions, where region A represents the surfaee for which the CFD model was developed. Region B represents the surfaee of the spike that exists in the same axial range as the nozzle liner. Regions A and B represents nozzle hardware extending from the inlet to the throat. Regions C and D represents the interior surfaees of the hardware that extends throughout the ehamber. The last region (E) eompletes the spike hardware from the throat to the tip.

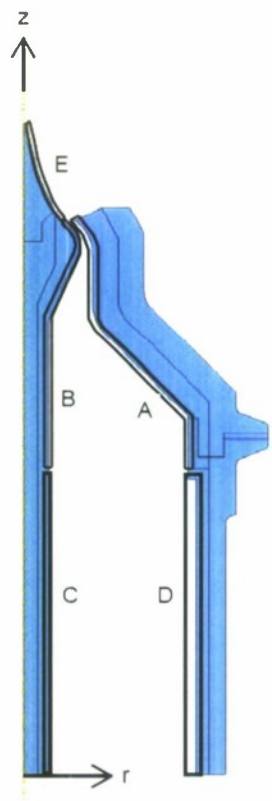


Figure 13.1 Diagram showing inner surfaees split into 5 regions for applying assumptions.

The first assumption allows for the direet applieation of the film eoeffieients at region A to region B. This assumption is supported using work earried out by Chung and Sung in 2003 that foeused on simulating turbulent flow through an annular pipe[30]. The size of the annulus for their study was eharaeterized by the ratio of the inner radius of the

pipe to the outer radius. The flow conditions were defined using a Re_D = 8900 and Pr = 0.71 for two radius ratios of 0.1 and 0.5. The same heat flux was assumed for both the inner and outer walls. The results of this study suggest that Nusselt numbers at the inner and outer surfaces differ by less than 20% for the 0.5 radius ratio, with the larger value being found at the inner surface. For a radius ratio of 0.1 the average difference grows to about 56% [30].

The proposed assumption is considered reasonable with these results in mind, in spite of the low Reynolds on whieh the study is based. The lowest radius ratio for the annular aerospike nozzle is 0.15. So, at the very least this assumption would yield erroneous eonveetive heat transfer eoefficients that are underestimated by up to 50% for the spike surface at the inlet. Yet, the coefficients at the inlet are more than an order of magnitude smaller than those near the throat. So, the effeets of this potential 50% error would be insignificant by comparison. Near throat the assumption gains validity where stations 10 through 16 have radius ratios that are greater than 0.5 where the Nusselt numbers are in better agreement, with a ratio of 0.9 at the throat.

This assumption only affords the spike a set of conveetive heat transfer coefficients from the inlet up to the throat using the CFD study results since this is where the nozzle liner ends. To approximate the coefficients for the rest of thc spike from the throat to the tip (region E) Long's correlation was used. Thc use of this correlation is supported by the good agreement exhibitcd in Figure 11.4. Finally, for the ehamber liner and the portion of the spike surfaee that exists upstream of the nozzle (region C and D) a uniform film coefficient will be assumed based on that of the inlet of the nozzle liner.

The last assumption made in applying the results of the CFD study to the transient heat conduction model regard how the convective coefficients are varied with time. Through the definition of steps these coefficients will be applied in a piecewise manner. For the first part of the piecewise function the coefficients will be ramped linearly to full magnitude over the start time (Step 2), or ignition delay time. This time represents the portion of rocket motor operation from the instant of ignition to the instant at which 10% of the maximum chamber pressure is achieved [21]. During this start time propane is combusted to provide a source of ignition. At exactly what instance the propane ignition source is turned on during the start time is unclear, as it depends on human interaction. Furthermore, it is difficult to characterize how this ignition event contributes to the overall heat transfer rate. But, by prescribing the linear ramping scheme, a conservative approach has been taken in considering the contribution made by the combusted propane to the heat of the nozzle wall. A sample plot of the transient film condition applied at the throat is given in Figure 13.2. This plot has been overlaid onto a sample pressure time trace from "*Run #20*".

After this start time it is assumed that the combusted gases have reached an approximate steady state condition with regard to the composition of the products. It follows, then, that the gas properties and corresponding local Nusselt numbers would also reach steady state. The second part of the piecewise function (Step 3) holds these values at full magnitude for the duration of the simulation. Based on the pressure-time trace for *"Run #20"*, the start time(Step 2) is approximated as 3 seconds. The time during which the coefficients are held at a constant magnitude is defined as 10 seconds (Step 3). This gives a total simulated operating time of 13 seconds.

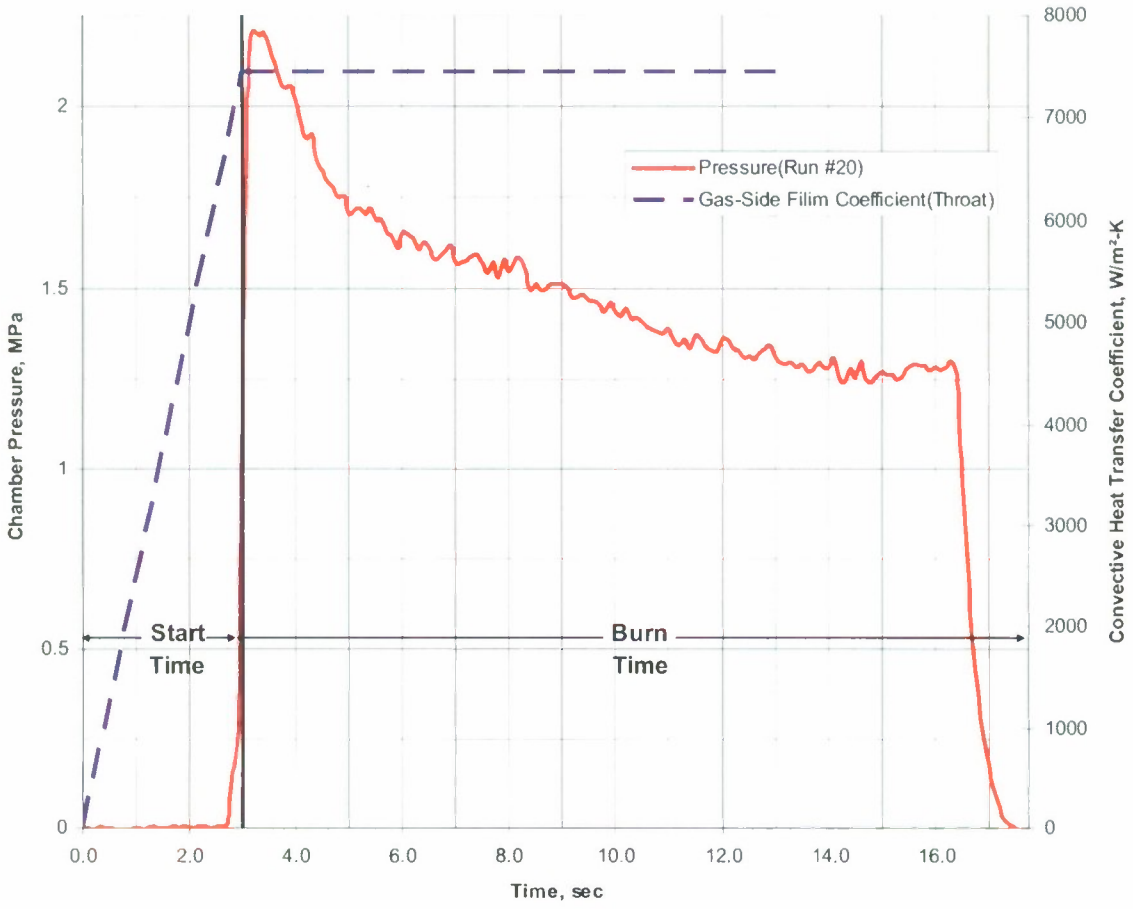


Figure 13.2 Plot of throat film coefficient as a funetion of time aeording to thc pieeewise seheme. Pressure-time traee overlaid for perseetive.

The results of the CFD model have been applied to a transient heat eonduction model developed in Abaqus. Assumptions were been made regarding the surfaces to whieh they can be applied and how their applieation is dependant on time. A summary of these boundary eonditions and the range of time during whieh they are applieable is listed in Table 13.1. The remainder of the development of the transient heat eonduction model involves defining the contact interaetions between components.

Table 13.1 Summary of boundary conditions

		Interactions (*Thermal Boundary Conditions*)			
		Surface to Surface Contact	Surface Film Condition		
Step	Simulated Time, sec.	Gap Conductance	Ambient Film Coeficient	Gas-Side Film Coefficient (*ramp to magnitude*)	Gas-Side Film Coefficient (*held at magnitude*)
initial	*t=0*	*Applied*	-	-	-
1		*Propagated*	*Applied*	-	-
2	*0<t<3*	*Propagated*	*Propagated*	*Applied*	-
3	*3<t<13*	*Propagated*	*Propagated*	*Propagated*	*Applied*

Surfaee-to-surfaee interaetions were ereated at the interfaee of every component in the assembly to define a gap conductance across mating surfaces. These interactions are applied in the initial step and propagated through the steps that follow. This gap conductance plays an important role in predicting how heat is conducted between two materials that are in contaet. The simplest approaeh to defining this interaction would be to assume that the surfaces are in perfeet eontaet with eaeh other so that there is no thermal resistanee. This assumption would imply that these surfaees are perfectly smooth. In reality machined surfaces have some measurable roughness. The effeet of this roughness is that only a portion of the mating surfaee areas are aetually in contaet. For the remainder of the mating surfaee area tiny voids exist between the surfaees and are filled with air. The air trapped in these voids acts like an insulator making the thermal

resistanee a non-zero value. The effective thermal resistanee used for this model is a funetion of the surfaee roughness of the mating eomponents and the properties of the air that lie in the voids ereated by that roughness [31, 32]:

$$h_{gap} = \frac{k_{gap}}{R_a + M} \tag{13.1}$$

Where k_{gap} is the gas thermal eonduetivity of air, R_a is the average surfaee roughness of the mating components and M is given by:

$$M = \alpha\beta\Lambda \tag{13.2}$$

Where α=2.4, β=1.7, and Λ= 0.06 μm. The α and β terms are dependent upon the properties of air and Λ is the mean free path of the air moleeules. The thermal conduetivity of air is approximated as 0.131 W/m-K for a temperature of 2100 K (3320 °F) [33]. The surface roughness was taken as an average of typical values for a boring or turning produetion method [34]. Using an average surface roughness value of 3.3μm, the gap conduetance was approximated as 37000 W/m^2-K. This gap conductance was applied uniformly at every surfaee-to-surface contact and was held constant for the entire simulated run.

The contact interactions for surface film conditions and surface to surface contacts have been defined for the transient heat eonduction model. These interactions were applied to a series of steps to simulate a "hot-fire" run for the annular aerospike nozzle. The resulting temperature fields ean now be mapped into the thermal struetural model to study how the various materials eombinations behave mechanically.

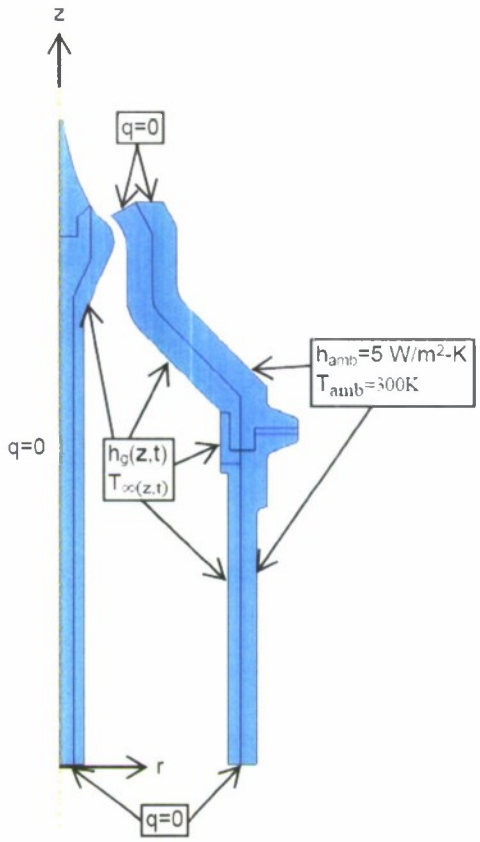


Figure 13.3 Diagram showing surfaces to which the ambient and gas-side film coefficients are applied for steps 2 and 3.

CHAPTER 14

Thermal Structural Model

The thermal structural model was defined by applying mechanical interactions between the various components, specifying boundary conditions to constrain the assembly and simulating loading cases. Tangential and normal behavior were specified for the interaction between components. A general coefficient of friction of 0.2 defined the tangential behavior and a "hard contact" with separation was defined for the normal behavior. This normal behavior definition will recognize when two components make contact and allow that contact to affect the mechanical stresses in each.

Boundary conditions were defined for the bottom surfaces of the spike, chamber liner, and chamber outer (Figure 14.1). These boundary conditions essentially prescribe a pinned joint, constraining the axial and radial displacements. In addition, the application of the pinned joint along these surfaces imposes a constraint on circumferential displacements. The axisymmetric geometry imposes a constraint on rotations about the radial and axial coordinates plus and a displacement in the circumferential direction.

Loading cases defined for this model include the effects of temperature gradients through thickness of each component and the Marman clamp that fixes the nozzle assembly to the chamber assembly. Mechanical loading due the distribution of static pressures throughout the nozzle was initially considered. But, early results revealed that these pressures did not contribute significantly to the stress response of the hardware. The inclusion of these pressures was abandoned to simplify the model. The inclusion of the Marman clamp is important because it captures initial conditions of the rocket motor before it is exposed to the extreme condition created by the hot gases. In some ways, the

Marman clamp can be considered a boundary condition that ties the nozzle outer to the chamber outer. In fact, this is partly true with regard to the way in which the clamp was modeled. The pressure that is applied by the Marman clamp to the angled surfaces of the nozzle outer and chamber outer is approximated as 29.2MPa. Sample calculations supporting this approximation can be found in Appendix H. Simply applying this pressure was found to be insufficient to fully constrain the model. To overcome this, a tie constraint was imposed to fix the gasket to the respective mating surfaces of the nozzle outer and chamber outer.

To predict thermal and mechanical stresses in the assembly, nodal temperatures are read into the model from the transient heat conduction results. Once the temperatures are applied to their respective nodes, displacements are created in each component based on the thermal expansion properties of the prescribed material. These displacements are restricted either directly or indirectly through the boundary conditions. These restrictions cause thermal stresses within each component as well as mechanical stresses from component interactions.

The nodal temperatures are read in through a predefined field load. The change in the temperature field with respect to time is preserved by setting up this predefined field to call on specific time steps from the transient heat conduction model. The thermal structural model is made dynamic in this way. In the definition of this field, the nodal temperatures for the heat transfer model are read in for every time increment in Steps 2 and 3 of the transient heat conduction model. This means the nodal stresses are calculated for the temperature field at every increment for Steps 2 and 3. The result is a variation of the resulting stresses with time. Since plastic hardening properties have been

specified for some materials the non-linear geometry option was turned on for steps 2 and 3 where yielding is expected in some of the isotropic materials. These time-dependent stresses are used to determine a maximum suggested operating time for the un-cooled annular aerospike nozzle.

Table 14.1 Summary of applied interactions, boundary conditions and loads

		Interactions		Boundary conditions		Loads		
		Surface to Surface Contact	Tie Constraint	$U_r=U_z=0$	$U_r=U_z=0$	Surface Pressure	Predefined Temperature Field	
Step	Simulated Time, sec.	Tangential and Normal Behavior	Gasket to Nozzle Outer and Chamber Outer	Bottom Surface of Spike	Bottom Surface of Chamber Liner and Outer	Marman Clamp	Nodal Temperatures, Step 2*	Nodal Temperatures, Step 3*
initial	*t=0*	*Applied*	*Applied*	*Applied*	*Applied*	-	-	-
1		*Propagated*	*Propagated*	*Propagated*	*Propagated*	*Applied*	-	-
2	*0<t<3*	*Propagated*	*Propagated*	*Propagated*	*Propagated*	*Propagated*	*Applied*	-
3	*3<t<13*	*Propagated*	*Propagated*	*Propagated*	*Propagated*	*Propagated*	*Propagated*	*Applied*

** refers to steps in the transient heat conduction model*

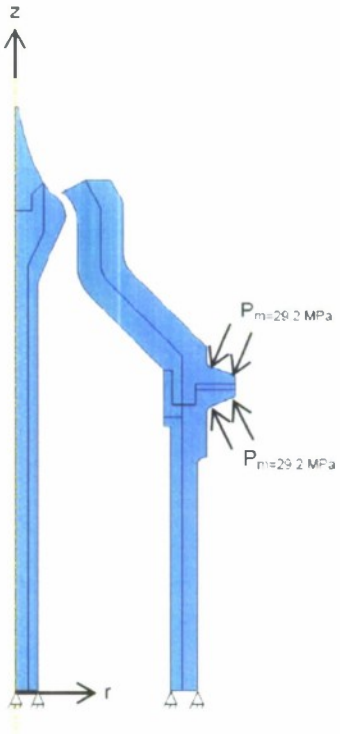


Figure 14.1 Diagram showing pin joint boundary conditions and Marman clamp load.

CHAPTER 15

Spike Web Model

In addition to thermal loading, the spike web is intended to be loaded in a direction that is normal to the axial coordinate. This intention is predicated on the knowledge that for an initial material choice of graphite the design goals for the spike web should not include any loading case that places the spokes of this structure in tension as a primary function. This is because graphite is a brittle material and inherently poor in tension. By limiting its sole function to providing lateral stability the design makes use of a compressive strength that is more than triple that of the tensile strength [9]. In application, the spike web is expected to experience very complicated dynamic loading. Unfortunately, the scope of this project does not permit a study of this magnitude. Instead, a static loading case is applied as a first step in the development of this component. The hope is that in making conservative assumptions about the static loading case the resulting design will be adequate for testing and a good starting point for future design iterations.

As has been mentioned before, a worst case scenario was imagined in order to load the spike web for analysis. The imagined loading case reduces the assembly of the spike and axial support rod to a simply supported beam where a distributed load is applied along a portion of the length of the assembly (Figure 15.1). A pinned support for the beam is prescribed at the bottom of the chamber (R_{Dy} and R_{Dz}) and a roller support represents the reaction at the inner hole of the spike web (R_{By}). The distributed load makes use of the static pressures within a rocket nozzle. A linearly varying load was defined using a chamber pressure of 3.10 MPa (450 psi) and a predicted static pressure of

0.155 MPa (22.5 psi) at the tip of the spike. The prediction of the static pressure at the tip was conducted using TEP and was based on a chamber pressure that is indicative of a brief peak pressure that occurs during operation. A linearly varying pressure function was defined based on the chamber pressure and static pressure at the spike tip. A corresponding linearly varying load distribution was derived from this pressure function. This was done by first determining the effective force acting on this area with regards to the distributed linearly varying pressure. This effective force was then applied along a portion of the beam extending from the spike tip to where the nozzle ends. Detailed calculations can be found in Appendix I. The reaction at R_{By} was found to be 2296 N (516.16 lbf).

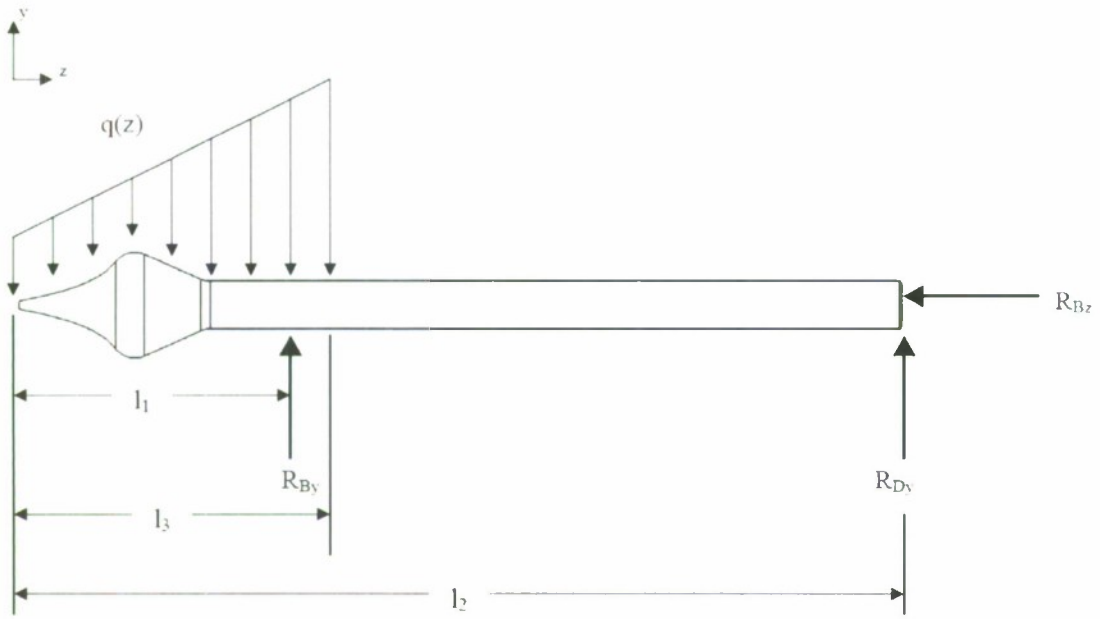


Figure 15.1 Loading diagram for spike.

Three concepts for the spike web were analyzed using the results from the above loading case. Simple FEA models were constructed using beam elements in Abaqus Standard. Beam elements are an ideal application for this analysis because they are

accurate, quick(computationally speaking) and work well with the simple geometry of the spike web. The spokes were represented by simple lines and the hub in the center was represented by a circle. Simple rectangular sections were used for the spokes and the hub. The thickness of the spokes was the only dimension that was varied across each model. Everything else was held constant, with the exception of the number of spokes in the third case. The outer rim was eliminated and characterized by fixed boundary conditions at the extreme end of each spoke. In addition, beams elements were added to the interior of the hub to simulate circular shape of the spike. If these extra beams were not applied the center would collapse under load and would not yield accurate results. The only drawback to using beam elements is that the do not capture stress concentrations where the spokes meet the hub. A more complex model was created later to address possible stress concentrations after first choosing a design using the beam element models.

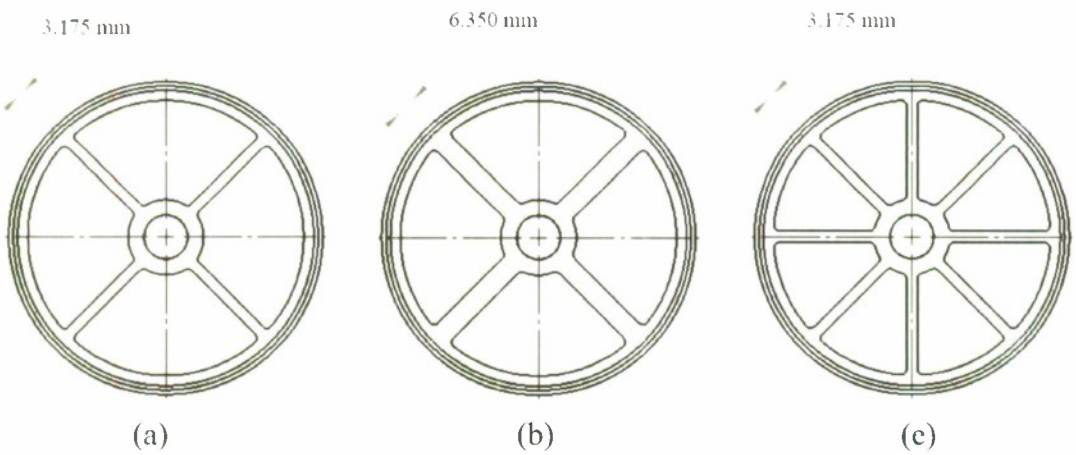


Figure 15.2 The three spike web concepts analyzed with web thicknesses dimensioned in units of meters.

Each model was analyzed by applying the 2296N load as a concentrated force in two directions. The first direetion applies the load along one spoke while the other applies the load direetly between two spikes. Figure 15.3 shows how the concepts were initially represented in Abaqus. Seetion forees and moments were extracted directly from Abaqus after each model was run. This information was then used to determine axial and bending stresses using elementary meehanies of materials. The maximum normal stress criterion was used to assess failure [35]. In addition, the largest eompressive force found for each of the three eoneepts was used to assess the potential for column buekling. The critical load used for this assessment was based conservatively on a pinned-pinned eonfiguration. A detailed analysis of the FEA results and column buckling assessment is given in Appendix J.

It is worth noting that the mere existence of the spike web at the inlet plane of the nozzle bloeks a portion of the eross seetional area and could adversely affect the performance of the rocket motor. This area can be eharaeterized by comparing the portion of the nozzle cross seetion that is blocked by the spike web to the area that would exist if no spike web was installed at all. The bloeked area is listed as percent of the nozzle cross-seetional area along with the results for this FEA study to offer some perspeetive on the potential performance trade-offs that have been made in providing strength to this component. It is reasonable to assume that the area blocked by the spike web will have very little effeet on the flow of the nozzle since the veloeities at the inlet of the subsonie region are typieally low.

Figure 15.3 FEA models used to initially analyze the spike web concepts.

Table 15.1 Summary of results for FEA models.

Concept #	*# of Spokes*	*Spoke Width, m (in)*	*% Blockage of Area*	*Largest Failure Index* (σ_1/σ_{ult})
1	4	0.003175 (.125)	10.2	0.906
2	4	0.00635 (.25)	17.7	0.5
3	8	0.003175 (.125)	17.8	0.46

The results for the spike web comparison study are listed in Table 15.1. Only the largest failure indiees are summarized, but a more eomplete listing ean be found in Appendix J. It should be noted that all three eoneepts yielded low failure indices with regard to the pinned-pinned eolumn buekling ealculation, eliminating this as a eoncern. The results show that eoneept #1 is the weakest design whieh makes sense because it uses only four spokes that have the smaller width speeifieation. This eoneept was included because it provides the smallest amount of bloeked area at 10.2%. Coneepts #2 and #3 eompare nieely with regard to their respeetive failure indiees and blockage area. The only advantage that eoncept #3 may have is that inereasing the number of spokes may inerease the components stiffness in a given direetion. Coneept #3 will be the better choiee simply beeause of the benefit of inereased stiffness.

To solidify the 8-spoked design a more eomplex FEA study was eonducted to see where stress eoneentrations exist. A detailed design was ereated in Solidworks that ineludes fillets where the spokes meet the hub and outer rim. A 2-D sketeh was pulled from this design and imported into Abaqus Standard. A plane stress model was created based on this sketeh and meshed with quadratie quadrilaterals. A brief grid density study suggested a global seeding size of 0.625 mm for the mesh. The FEA models used for the eomparison study were employed to help verify the inputs and setup for this plane stress model. The use of a eoneentrated foree was abandoned for this study to eliminate the

influence of unrealistic results at the point of application. Instead, a dummy cross section was created to represent the portion of the spike shank that would make contact with the hub of the spike web. A body force was applied to this component using the prescribed load of 2296 N. Pinned boundary conditions were applied to the outer rim of the geometry.

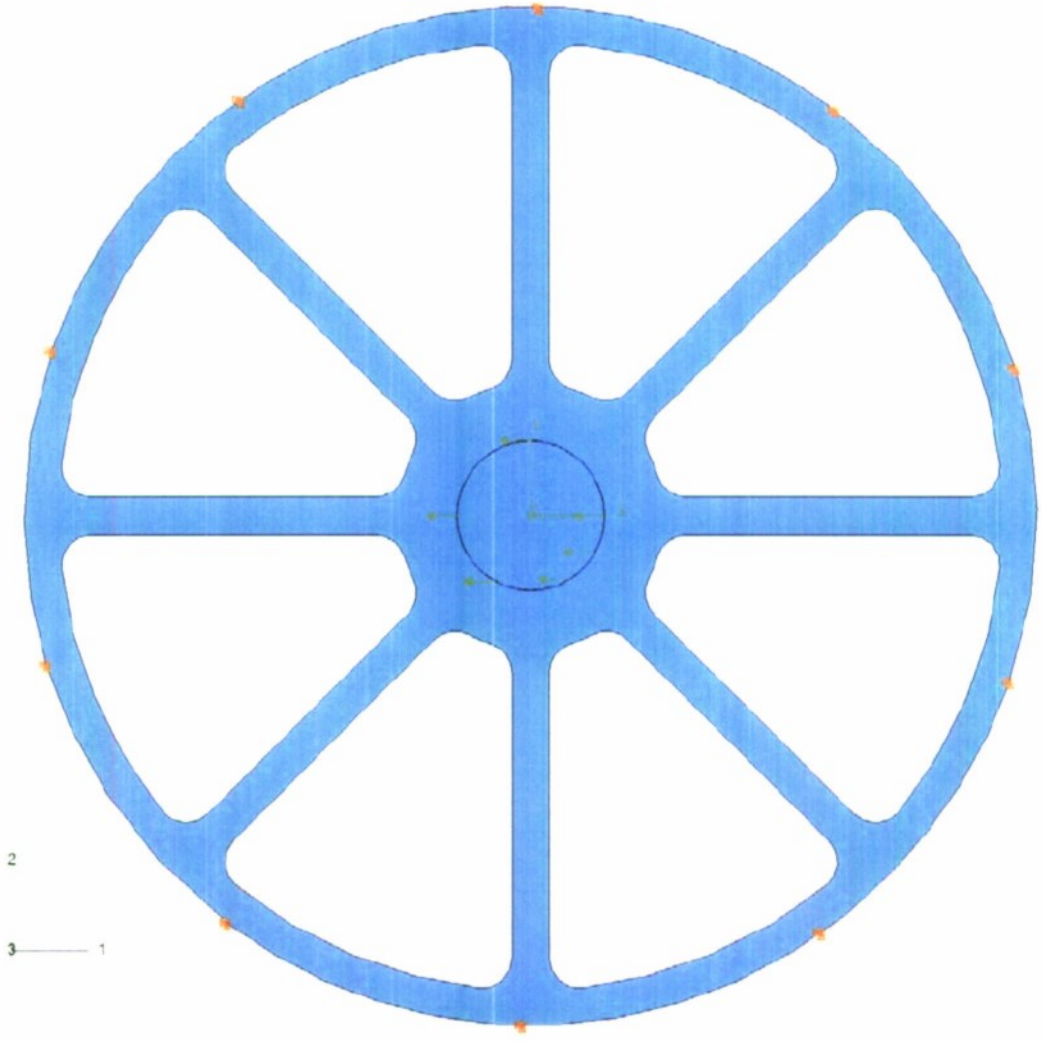


Figure 15.4 Example of the 2-D plane stress model geometry set up in Abaqus.

The final portion of the spike web analysis focuses on the thermal stresses generated by temperature gradients through this component. The approach is similar to

what was deseribed in Chapters 13 and 14. A transient heat conduetion model was set up to generate a temperature field that could be passed into a thermal structural model. The same time steps were used to simulate a test-fire. The main difference is that the spike web was modeled as a three dimensional geometry. However, the geometry was simplified to a sliee that represents 1/8 of the entire part and includes only one spoke to eut down computational expense.

Every surfaee of the spike was subjeeted to the film eoeffieient user-subroutine except for those where the model was eut and at the outermost and innermost eireumferential faces(Figure 15.6). The resulting time-dependent temperature field was then mapped into a thermal struetural model. In this model the spike is allowed to freely expand radially and axially, but symmetry boundary eonditions restriet displaeements at the eut surfaees eireumferentially. The maximum displaeements from this free expansion were eompared to the maximum displaeement at the surfaces where the spike web would mate with the nozzle liner. Displaeement data was obtained from the results of the assembly model to support this eomparison. The differenee in the radial displaeements for both of these eomponents was applied to the outermost surfaee of the spike web as a boundary eondition. A fit tolerane of the outer eireumferential surfaee of the spike web to the mating surface on the nozzle liner was speeified as .0508 mm (0.0020 in.) and ineluded in this differenee ealeulation. The resulting imposed displaeement value was approximated as 0.14 mm (0.0055 in.) inward in the radial direetion. In this way the radial foree exerted on the spike web by the nozzle assembly is simulated. The resulting stresses generated by this displaeement are then analyzed.

Figure 15.5 Section of a 3-D geometry of the spike web.

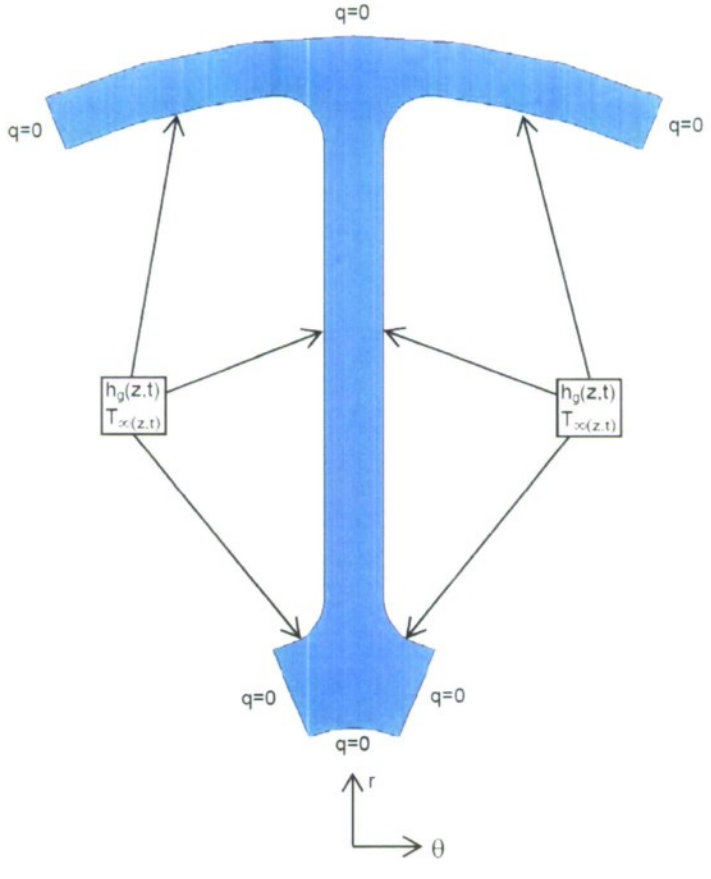


Figure 15.6 Thermal boundary conditions for the spike web.

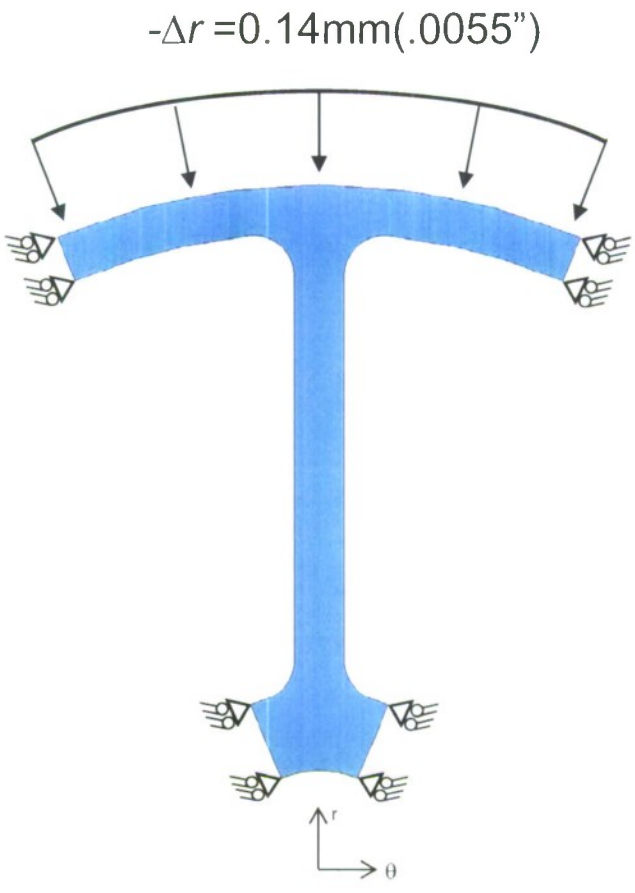


Figure 15.7 Static boundary conditions and imposed radial displacement for the spike web.

CHAPTER 16

Results and Discussion

Transient Heat Transfer Model

Eight different configurations in all were analyzed for the spike employing 5 different materials. Maximum allowable temperatures were defined for each material based on the highest temperature for which properties are available. Note that in the case of the coated SiC-6 grade graphite the maximum temperature is assumed to be the same as for the C/SiC. This assumption is founded on the idea that the graphite will erode significantly without the aide of a SiC coating. It is believed that this SiC coating should be subject to the same operational limits as the composite that uses the same material in its matrix.

Table 16.1 Maximum temperature allowed for each material[8].

Material	Design Allowable Temperature	
	K	°F
C/SiC	2273	3632
Inconel 625	1144	1600
Haynes 188	1255	1800
Columbium	1700	2600
Coated SiC-6 Graphite	2273	3632

The specific configurations that were analyzed and the corresponding results are given Table 16.2. The listed times were determined by identifying the instance during the simulation when any one of the materials used exceeded the design allowable temperature. It should be noted that the operational times listed represent an overall total including the start time and the burn time. Recall that the start time is given as 3 seconds and the burn time is 10 seconds.

Table 16.2 Results for total allowed time of operation. Total time means the start time plus the burn time.

Configuration #	Spike Outer Sleeve Material	Spike Inner Sleeve/Spike Tip Material	Predicted Total Operational Times, sec.
1	C/SiC	C/SiC	4.3
2	C/SiC	Inconel 625	3.1
3	C/SiC	Haynes 188	3.3
4	C/SiC	Columbium	3.6
5	Graphite	Inconel 625	2.9
6	Graphite	Haynes 188	3.1
7	Graphite	Columbium	3.7
8	Columbium	Columbium	3.1

The results are not very promising, suggesting that there is no material combination that results in an effective solution for a spike of this scale. The best configuration is made up entirely of C/SiC, which is estimated to last only 4.3 seconds after ignition. The top three performers were examined more closely to pinpoint exactly what component and/or material is the limiting factor in each combination. Detailed plots of temperature versus time for these three configurations are given in Figures 16.1, 16.3 and 16.5. Supporting contour plots are given in Figures 16.2, 16.4 and 16.6. These contour plots correspond to the time called out on the temperature versus time plots.

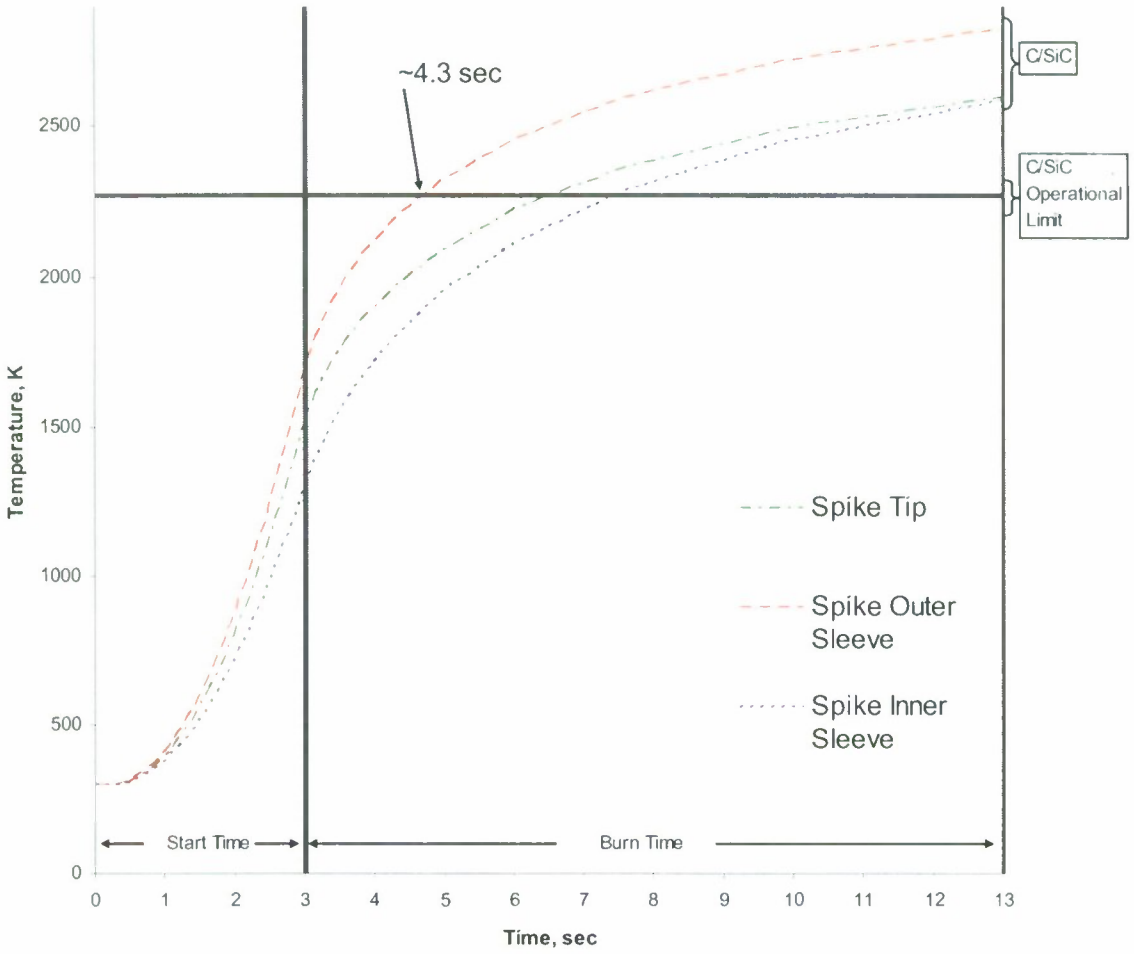


Figure 16.1 Temperature time trace of an all C/SiC construction.

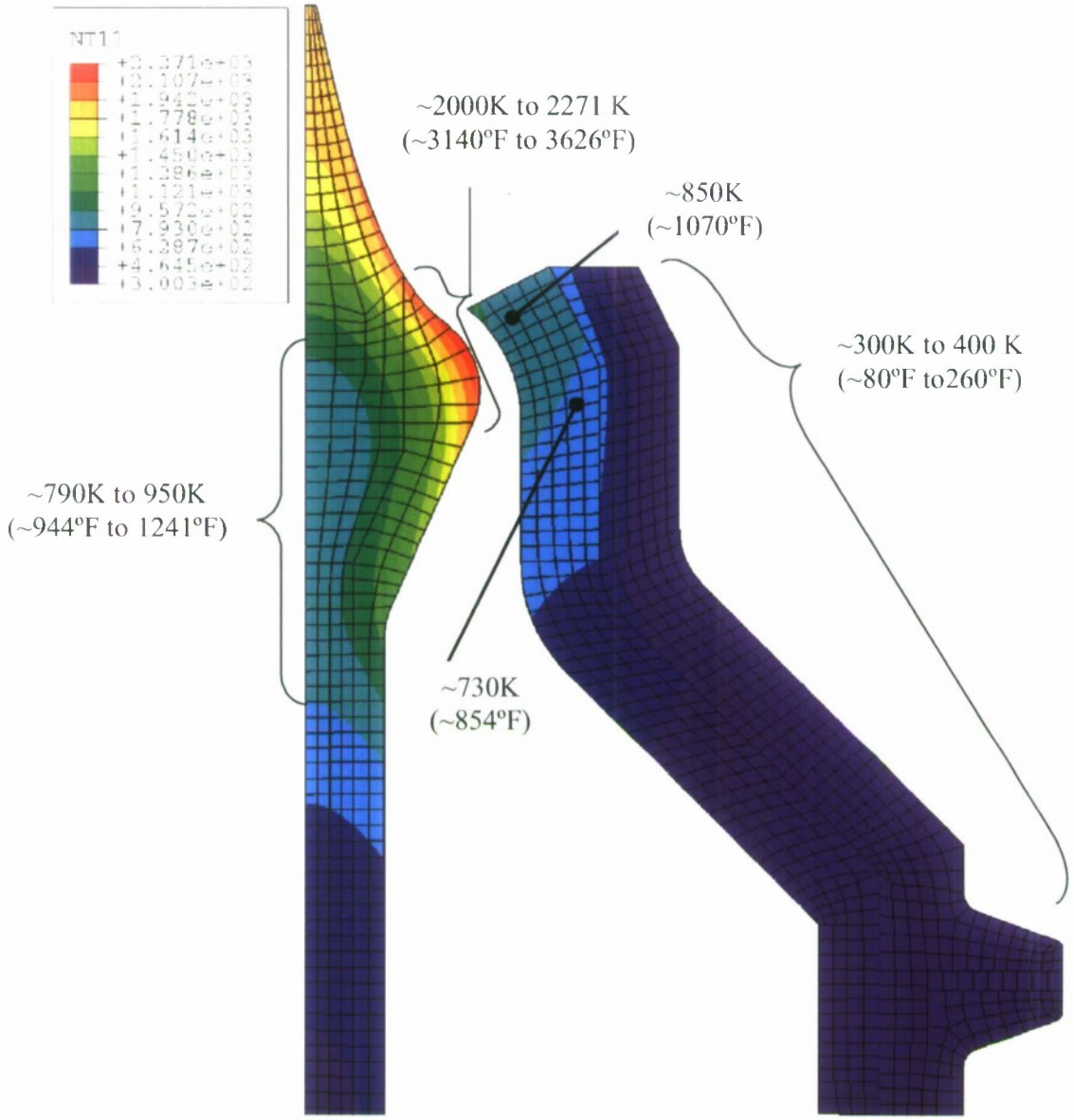


Figure 16.2 Contour plot of temperature at 4.3 seconds for configuration #1. NT11 is temperature in units of Kelvin.

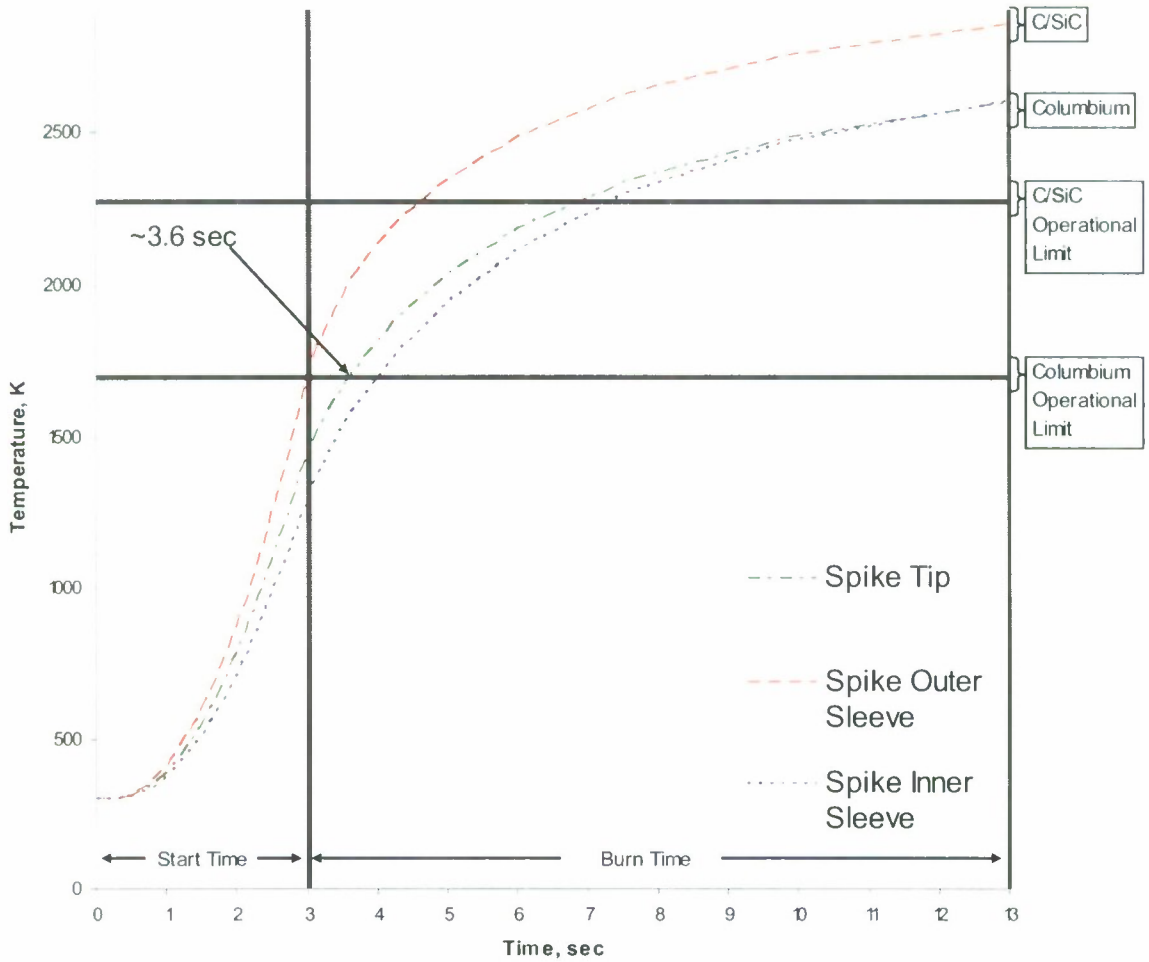


Figure 16.3 Temperature time trace of a part C/SiC and part Columbium construction.

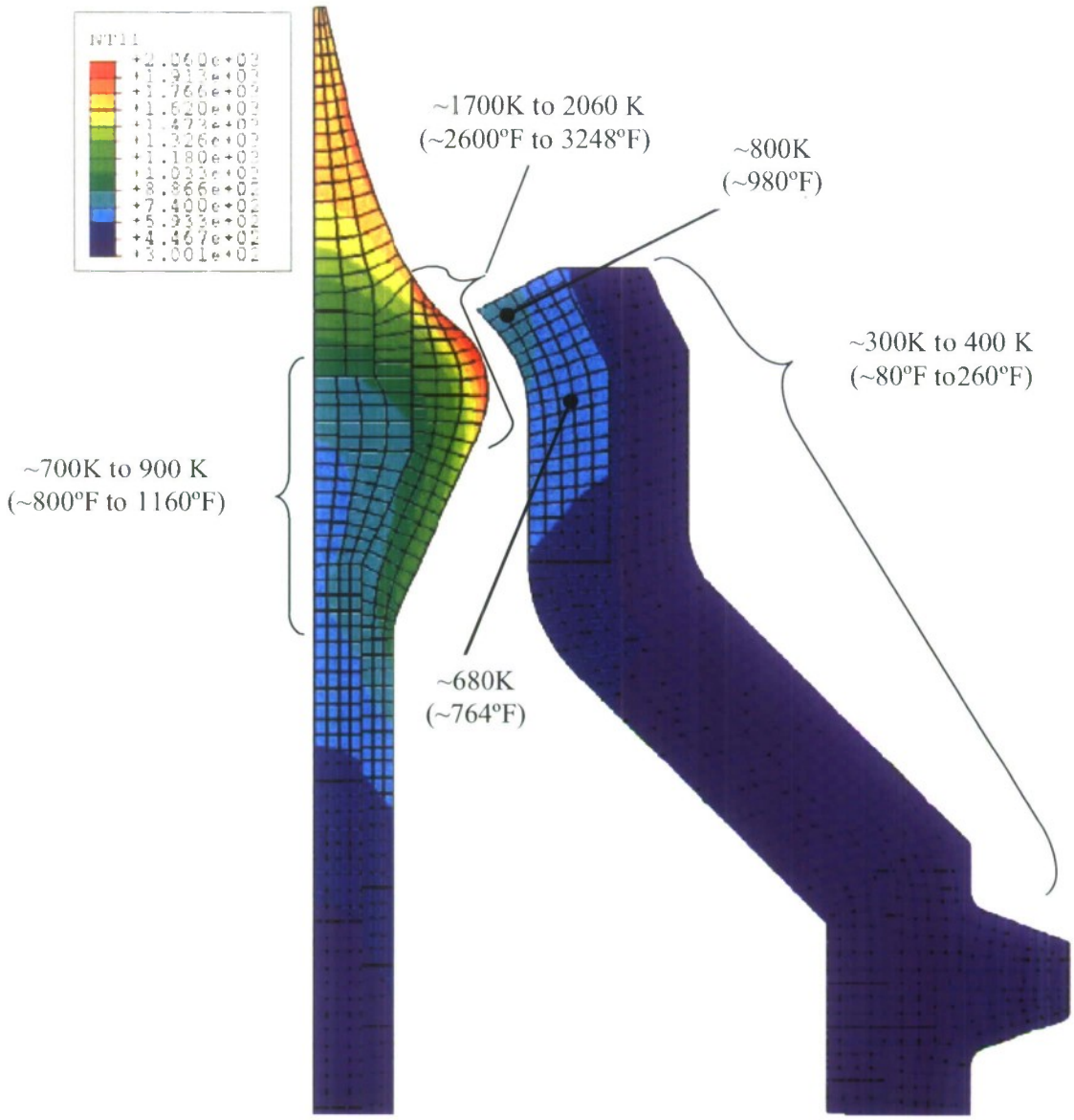


Figure 16.4 Contour plot of temperature at 3.6 seconds for configuration #4. NT11 is temperature in units of Kelvin.

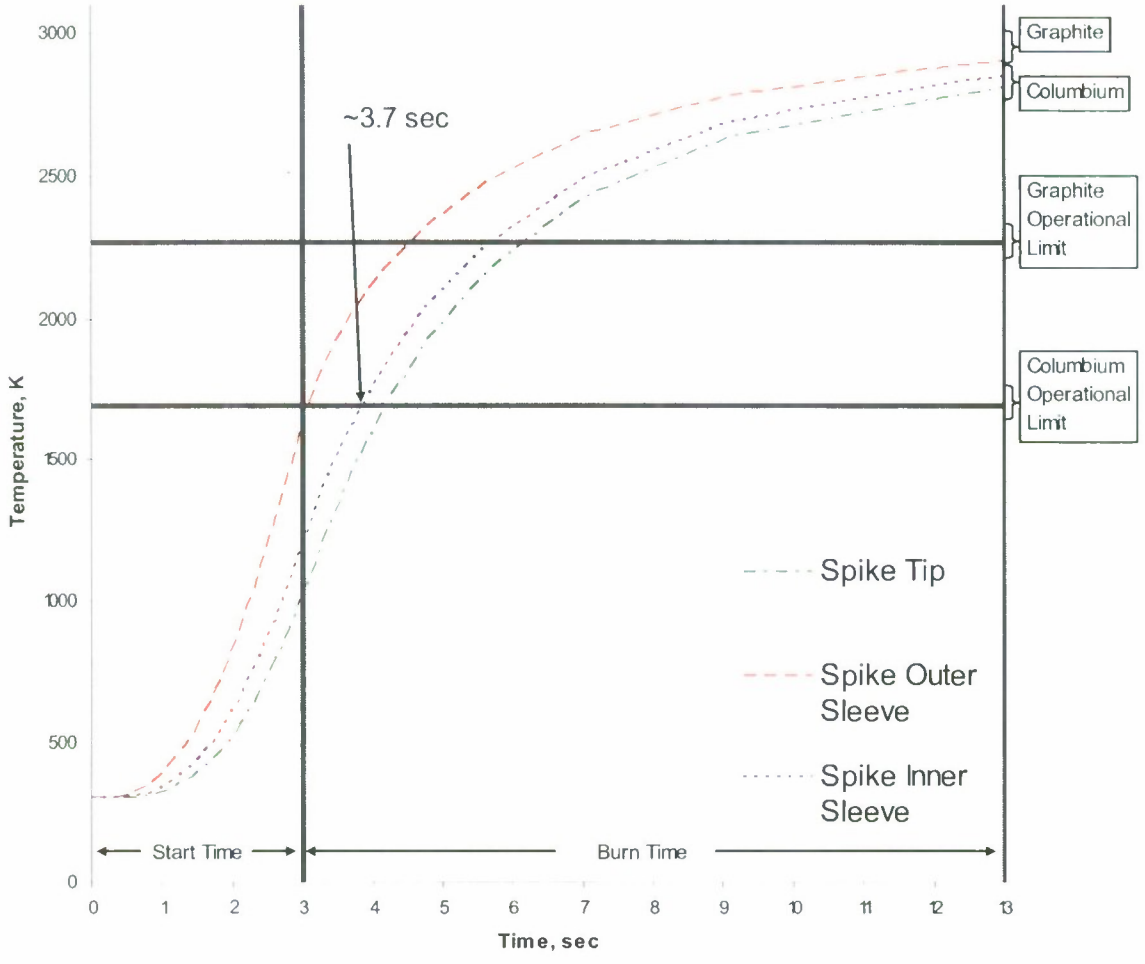


Figure 16.5 Temperature time traee of a part Graphite and part Columbium construction.

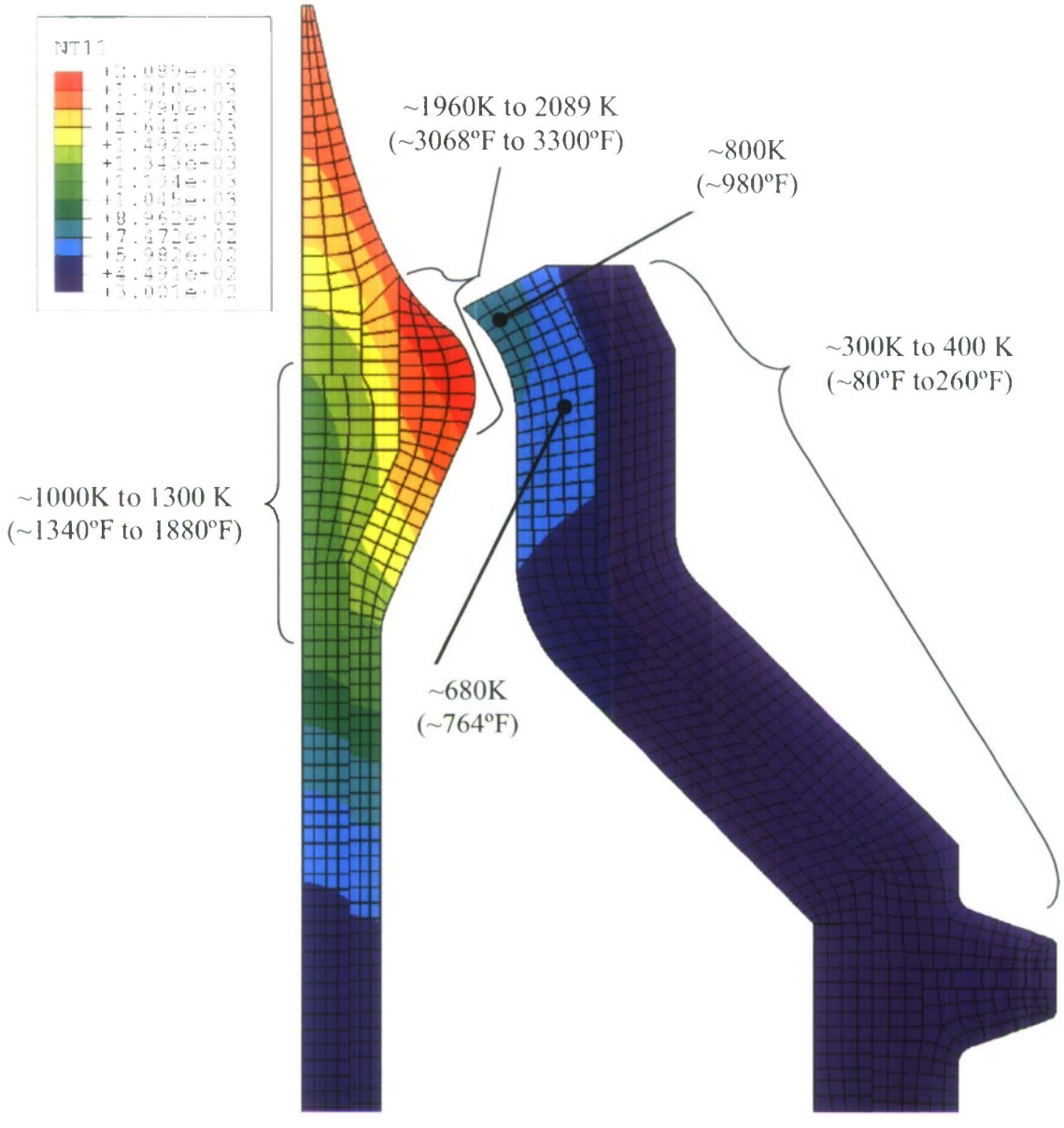


Figure 16.6 Contour plot of temperature at 3.7 seconds for configuration #7. NT11 is temperature in units of Kelvin.

These configurations were analyzed further in the thermal structural model. The results presented here reflect nodal locations at which the temperature was the highest for the respective spike subcomponent. However, the thermal stresses at these nodal

locations did not neeessarily reflect corresponding maxima. In most cases a different set of nodal locations was chosen that refleeted maximum thermal stress values.

Thermal structural Model

Results from the thermal stress model are organized in a similar way to the transient heat eonduetion model. The inerease in stresses due to thermal expansion and the mechanieal loading eaused by eomponent to component eontact is plotted versus time. These stresses are plotted along with ultimate strengths as a way to show where the subeomponent of the spike would likely experienee failure. For a material selection a Columbium FS85 a plot of yield strength is ineluded to prediet a time to yield as well as a time to failure. Temperature dependent yield and ultimate strength data for Columbium FS85 were approximated from curves in a specialized design allowables handbook[8]. This data was made dependent on time by eorrelating it to how the temperature at the given node of interest changed over the start time and burn time. In this way, time-dependent design allowables were ereated for Columbium FS85 with regard to a specific node. Unfortunately, similar data was not available for the C/SiC eomposite or the SiC-6 grade graphite, therefore ultimate strength properties were made independent of time and temperature.

The failure criterion used in analyzing eaeh of the subeomponents of the spike varied depending on the material used. For the Columbium alloy a Von Mises failure eriterion was used since the material is isotropie. The maximum normal stress eriterion was used for Graphite sinee it is a brittle material. The same failure eriterion was applied to the C/SiC eomposite. In reality, failure analysis for a eomposite is eomplicated and typieally addressed with a Tsai-Hill or Tsai-Wu eriterion [14]. However, for this

preliminary design effort a maximum stress failure criteria was deemed sufficient for an early look at using C/SiC.

In Figures 16.7 and 16.8 plots of stress versus time are displayed for the three subcomponents of the spike. The material configuration considered is one that employs Columbium FS85 for the spike tip and inner sleeve and SiC-6 grade graphite for the outer sleeve(configuration #7). The results suggest predicted operating times that are less than the initial predictions made based on the results of material temperature alone. These results are rather discouraging because they imply that this spike design would likely not last through the start time. Even more time is lost if a factor of safety is applied. The limiting factor in this design is the Columbium alloy employed for the spike inner sleeve. This subcomponent is responsible for taking axial loads and so failure could result in the throat being plugged. Plugging the throat of a rocket nozzle is considered catastrophic if no provisions exist that would allow the chamber pressure to be relieved through an alternate orifice.

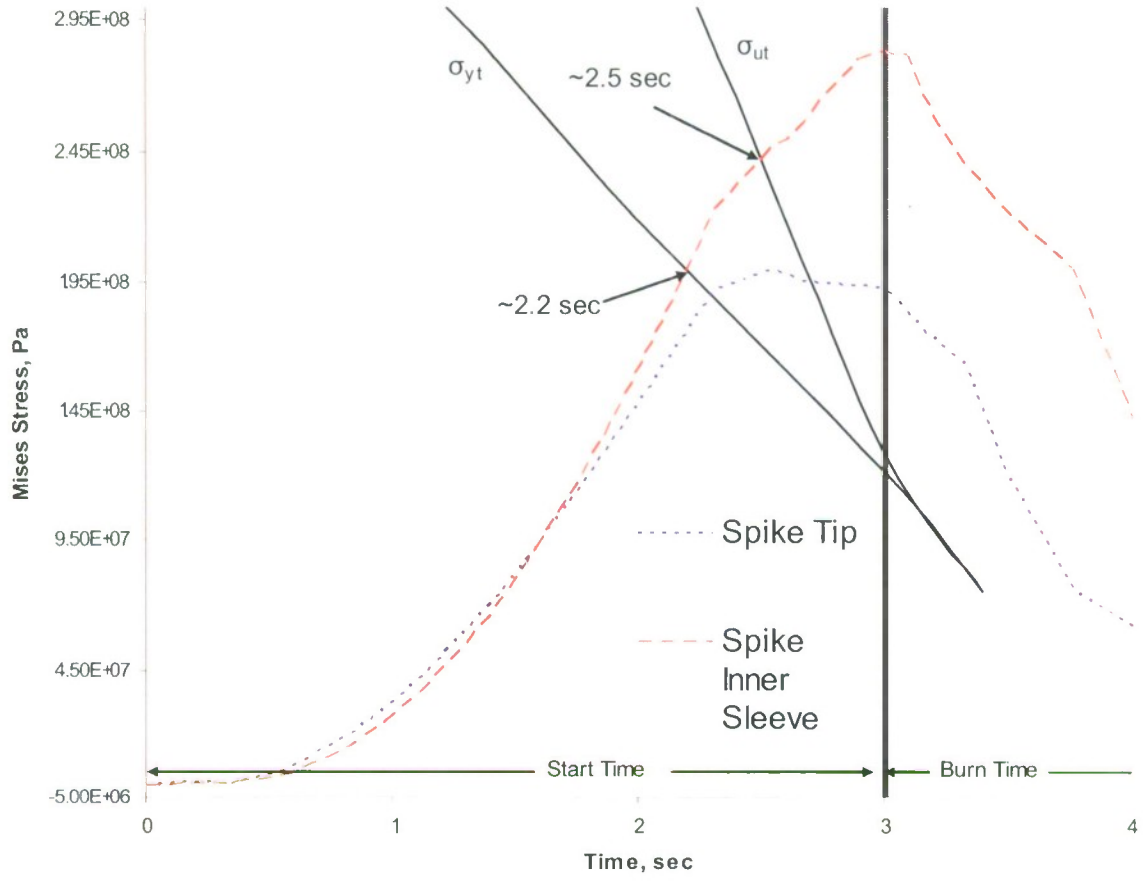


Figure 16.7 Von Mises stress versus time for Columbium employed at the tip and inner sleeve. Here graphite is used for the outer sleeve. Plot has been truncated to 4 seconds.

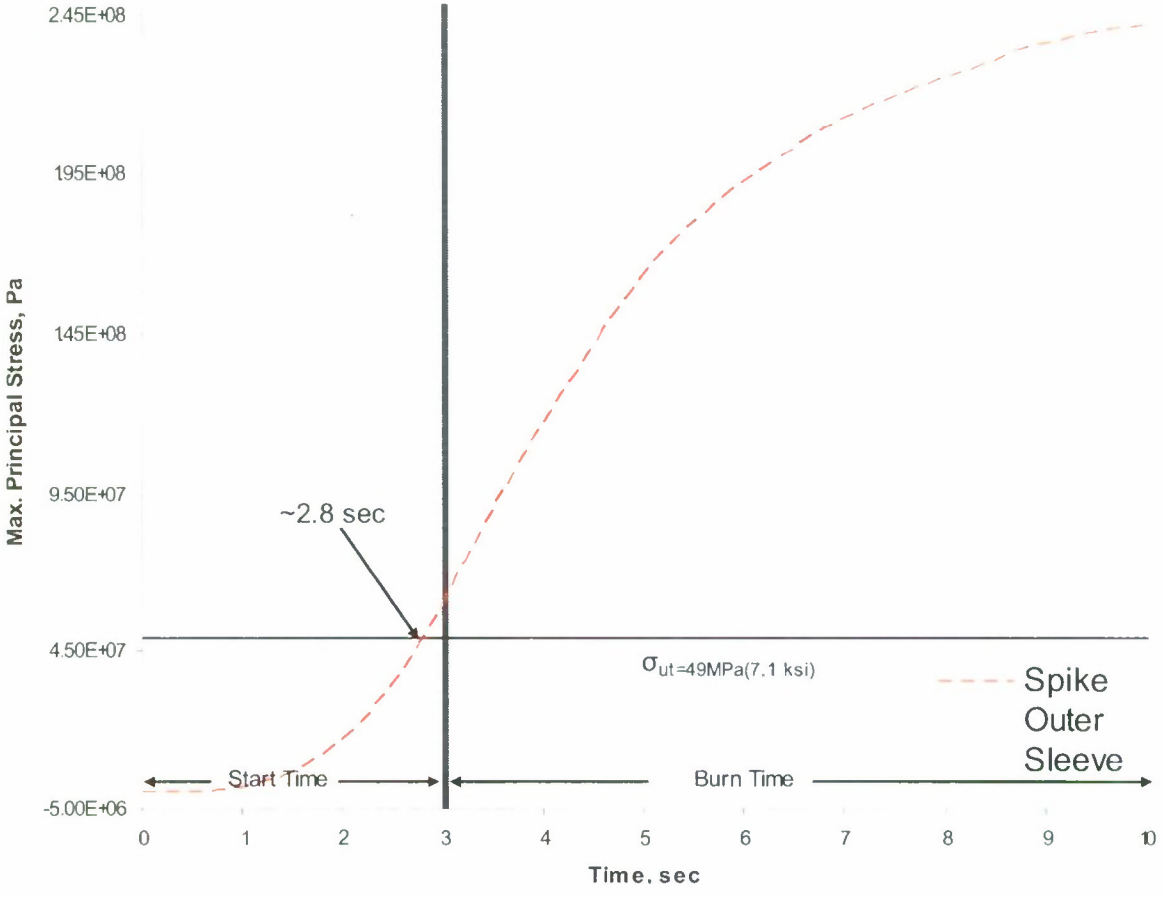


Figure 16.8 Maximum principal stress versus time for graphite employed for the outer sleeve. Plot has been truncated to 10 seconds.

Figures 16.9-16.11 display results for a material configuration that uses the Columbium alloy for the spike inner sleeve and spike tip and C/SiC for the spike outer sleeve (configuration #4). The thermal response of the Columbium is similar to Figure 16.7. Results for the C/SiC require some background before any discussion can proceed.

Recall that the material properties were applied in a manner that does not represent actual hardware (Chapter 12). The effect of this essentially changes the modeled fiber orientation of the composite so that instead of being orthogonal in three principal directions with respect to a cartesian coordinate system the orthogonality is now

relative to a cylindrical coordinate system. This means that the out of plane properties were modeled to represent a series of circumferential fiber "hoops". The highest stress in this direction that exists in the spike outer sleeve is plotted against time in Figure 16.10. The results in Figure 16.11 represent the largest stress among both of the principal directions for the r-z plane.

The portion of the spike outer sleeve where this maximum tensile stress exists is at the interface with the Columbium spike inner sleeve near the throat. In addition, large compressive stresses were found at the gas-side surface, also near the throat. This is significant because it means that the outer surface is being heated up and wants to expand radially. However, this expansion is not uniform because the interior surfaces of the spike outer sleeve are at lower temperatures. In fact, the thermal gradient across the radial thickness of the spike outer sleeve is quite large. This large gradient is directly related to a relatively low thermal conductivity for the C/SiC of approximately 36 W/m-K at room temperature[16]. The effect of this large thermal gradient is a resistance by the interior surface to the thermal expansion of the gas-side surface. This resistance is manifested by a circumferential stress at the interior surface that is tensile and one at the exterior surface that is compressive.

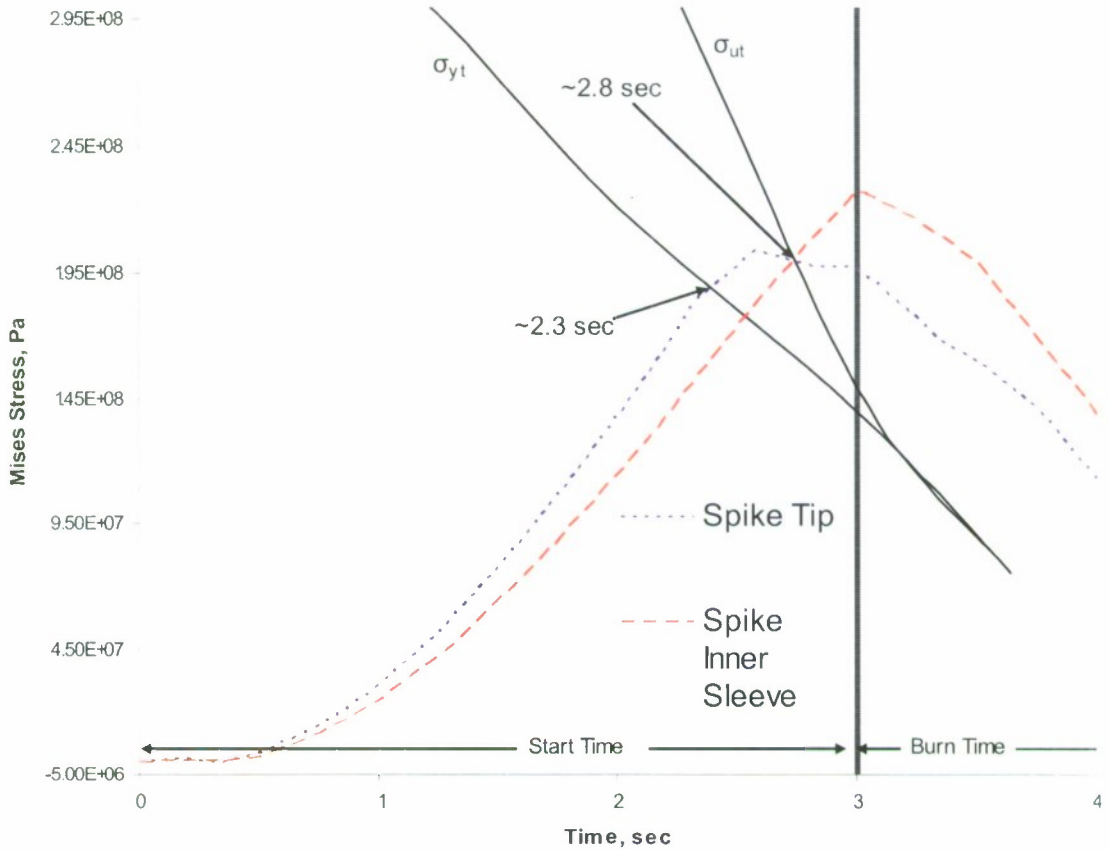


Figure 16.9 Von Mises stress versus time for Columbium employed at the tip and inner sleeve. Here C/SiC is used for the outer sleeve. Plot has been truncated to 4 seconds.

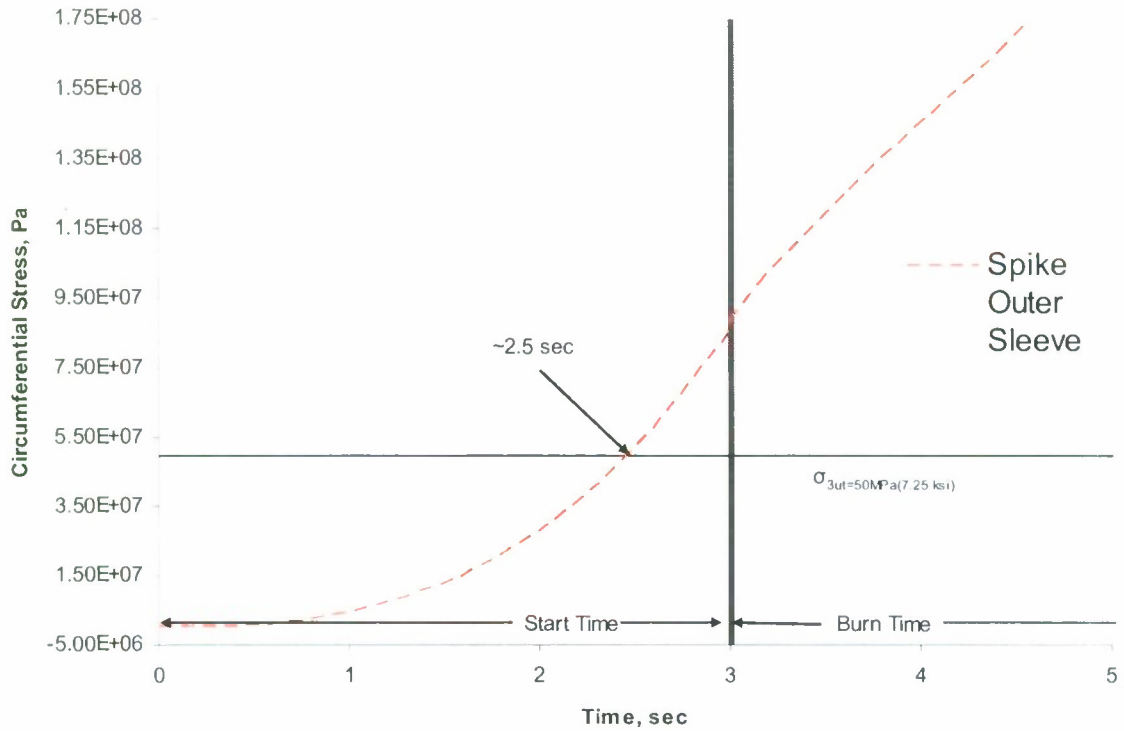


Figure 16.10 Circumferential stress versus time for C/SiC employed for the outer sleeve. Plot has been truncated to 5 seconds.

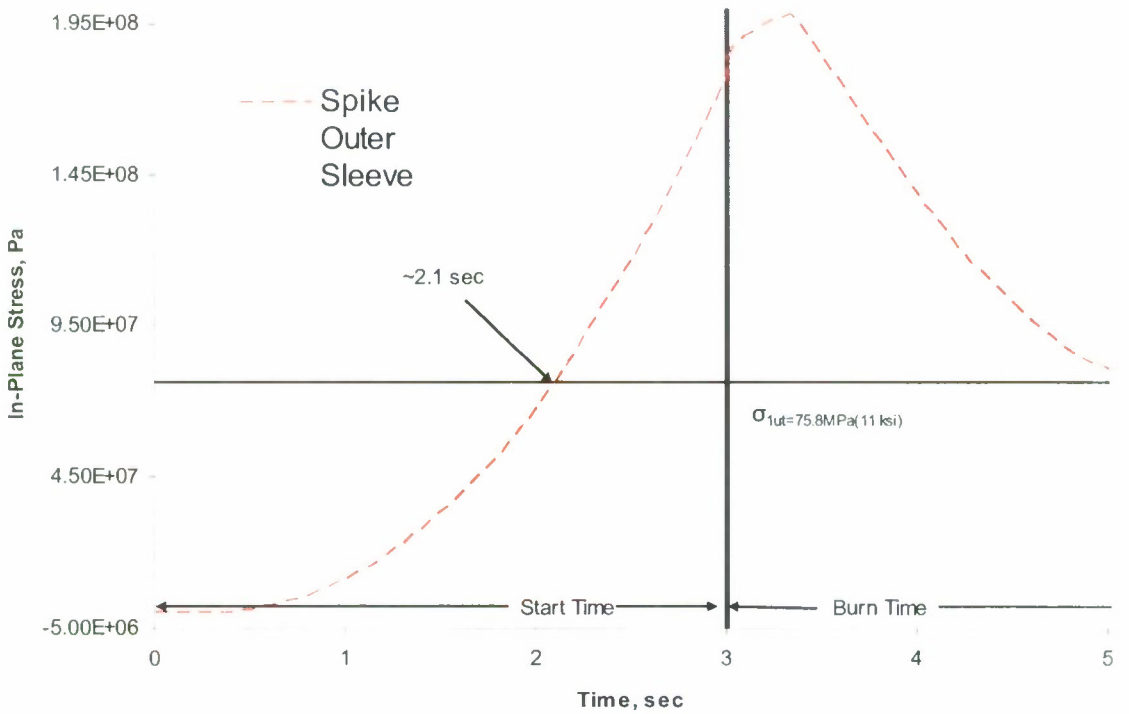


Figure 16.11 In-plane stress versus time for C/SiC employed for the outer sleeve. Plot has been truncated to 5 seconds.

For the results given in Figures 16.9-16.11 the properties in the material 1 and 2 directions (having the same properties) are aligned with the r-z plane. The out plane properties (modeled for the θ-direction) represent the weakest directions and also the softest, with regard to the material stiffness matrix. To put this in perspective, the as-modeled circumferential stiffness is significantly less compared to the other two material directions. To gain some perspective on how the "as manufactured" C/SiC material would behave in a different direction the stiffness properties were rotated 90° about the z-axis. This puts the weakest fiber direction aligned with the r-axis and maintains one of the stronger fiber directions aligned with the z-axis(axial coordinate direction). The out of plane material properties, now aligned with the θ-direction, are represented with one of the stronger fiber directions.

This material orientation has little effect on the subcomponent modeled with the Columbium alloy as can be seen in Figure 16.12. However, the response of the C/SiC changes significantly. The circumferential stress in the Figure 16.13 exceeds the ultimate strength of the C/SiC in the principal-1 fiber direction 4/10 of a second sooner. This is believed to be related to stiffer elastic properties that have been assigned to that direction. With or without the material rotation the C/SiC does not seem to perform as well as the graphite. Although the way in which the material was assigned to the spike outer sleeve subcomponent is not physically representative, some conclusions about how the actual hardware would perform in a hot-fire scenario can be drawn regarding the effects of a comparatively low thermal conductivity

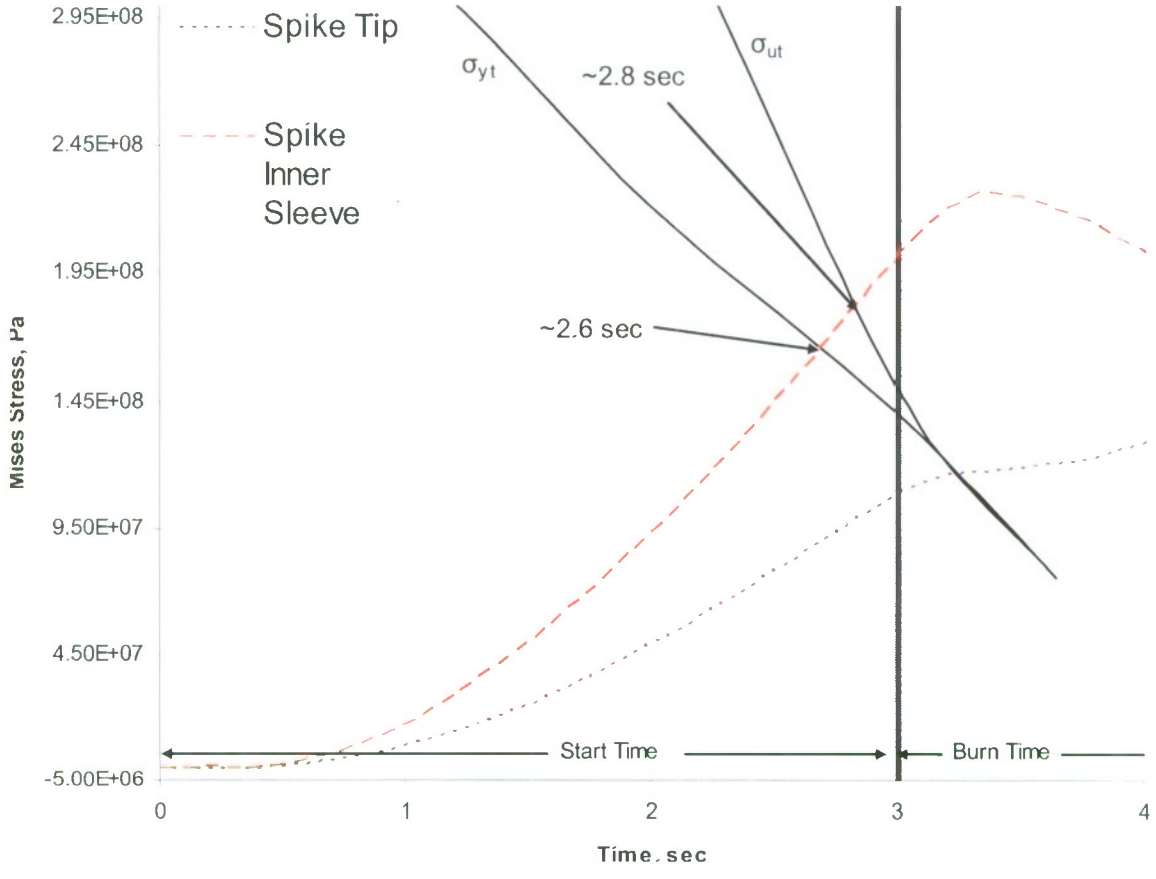


Figure 16.12 Von Mises stress versus time for Columbium employed at the tip and inner sleeve. Here C/SiC is used for the outer sleeve but with 90° material rotation about the axis of symmetry. Plot has been truncated to 4 seconds.

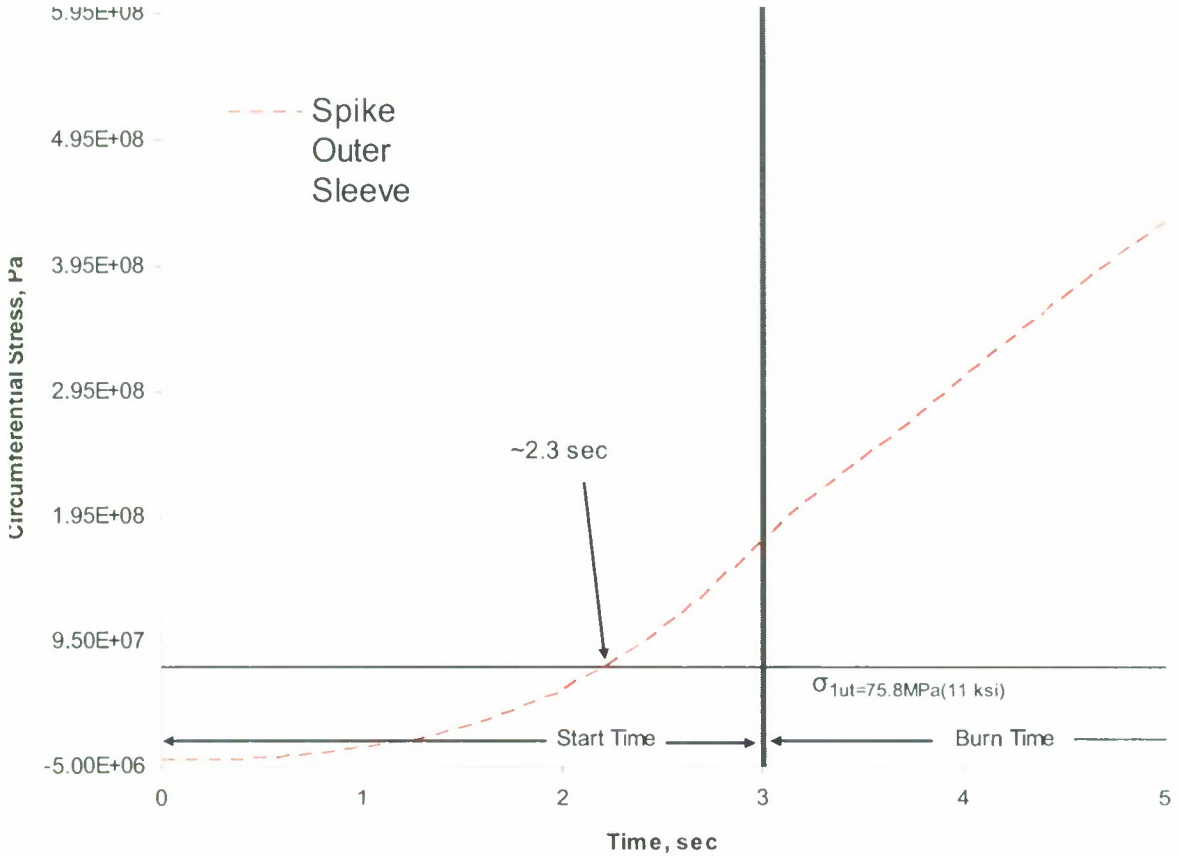


Figure 16.13 Circumferential stress versus time for C/SiC employed for the outer sleeve. Here C/SiC is used for the outer sleeve but with 90° material rotation about the axis of symmetry. Plot has been truncated to 5 seconds.

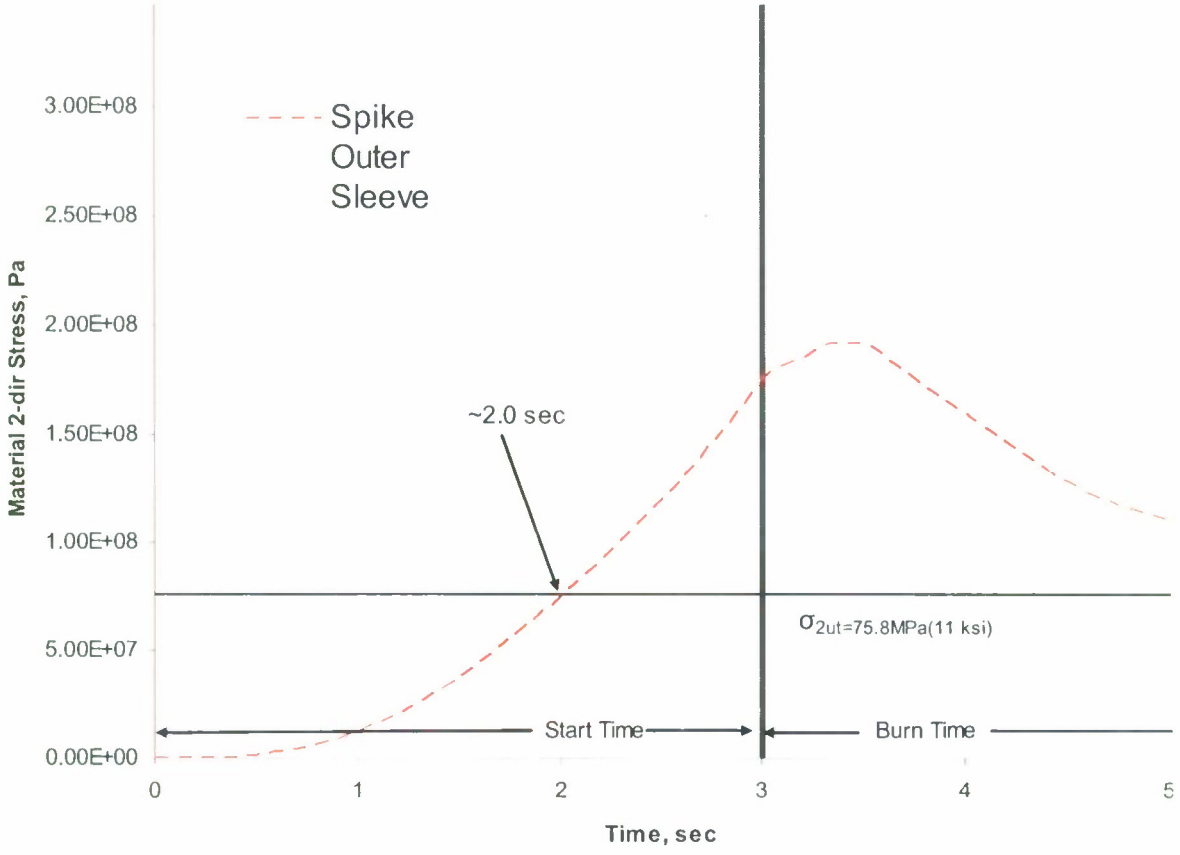


Figure 16.14 Principal 2-direction stress versus time for C/SiC employed for the outer sleeve but with 90° material rotation about the axis of symmetry. Plot has been truncated to 5 seconds.

In the final configuration for the spike C/SiC is applied to the entire part(configuration #1). The same conclusions about the 90° material rotation can be drawn and so those results are not displayed here. Such results would, in fact, be of little consequence since the configuration does not perform as well as the first two with no material rotation at all. Again, this can be attributed to the relatively low thermal conductivity of C/SiC. Large thermal gradients were also seen in this configuration, however the temperature difference is greater because the entire part is subject to the same thermal conductivity. To give an example, at 4.3 seconds(just after the simulated start time) the center of the spike near the throat is between 790 K (944 °F) and 950 K

(1241 °F) while the gas-side surface near the throat has increased to between 2000 K (3140 °F) and 2270 K (3626 °F). A contour plot of these temperatures is given in Figure 16.2.

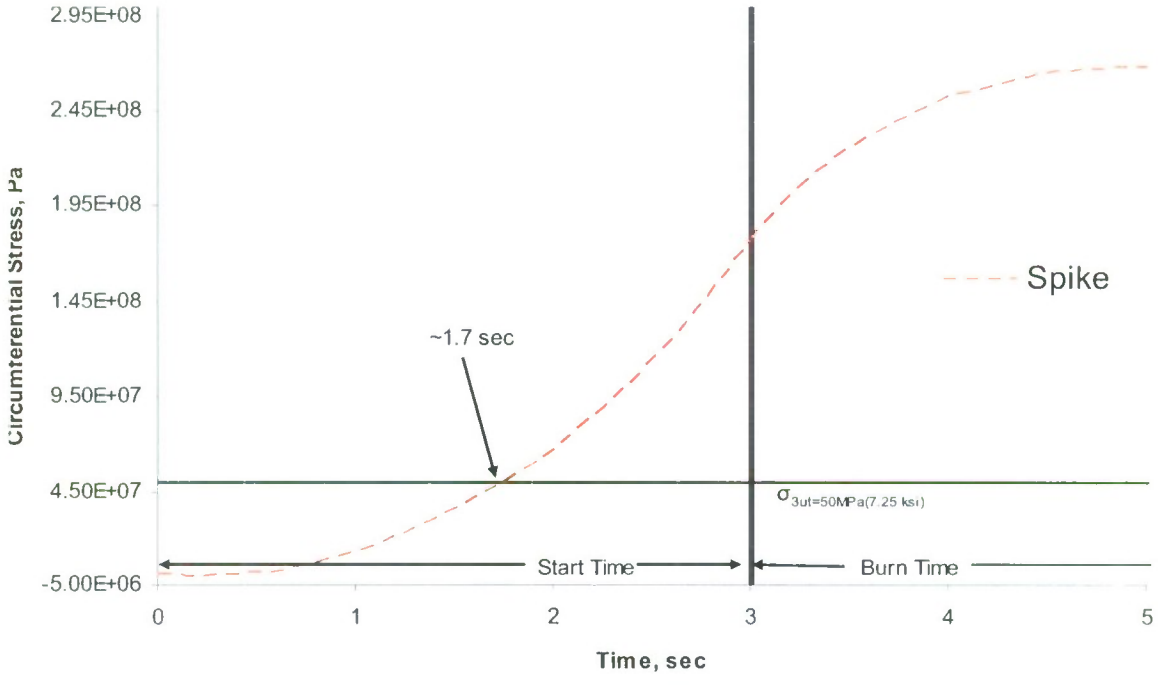


Figure 16.15 Circumferential stress versus time for C/SiC employed for the entire spike. Plot has been truncated to 5 seconds.

The results given for the various spike configurations are predicted to last a very short period of time before the first instance of failure. All of the configurations display approximately the same operating time and no one configuration has shown favorable results compared to the others.

Finally, in Figures 16.16 and 16.17 results are displayed for the nozzle outer and liner, respectively. The stainless steel nozzle outer experience's yielding after approximately 2 seconds, but never reaches the ultimate limit. The location of this maximum stress is on the ambient side of the part, radially adjacent to the throat. To

check the accuracy of this result a calculation was performed using temperatures at the inner and outer surfaces of the nozzle outer component. An analytical solution for a long cylinder with a concentric hole was used to approximate the circumferential stresses at the exterior [36]. The analytical solution showed good agreement considering that it uses a logarithmic temperature distribution and the actual hardware only locally represents a cylinder. This not only verified the reasonable accuracy of the results from Abaqus, but also revealed that the largest contribution to the Von Mises stress is from the circumferential stress. This effect can again be attributed to the thermal gradient across the thickness of the part.

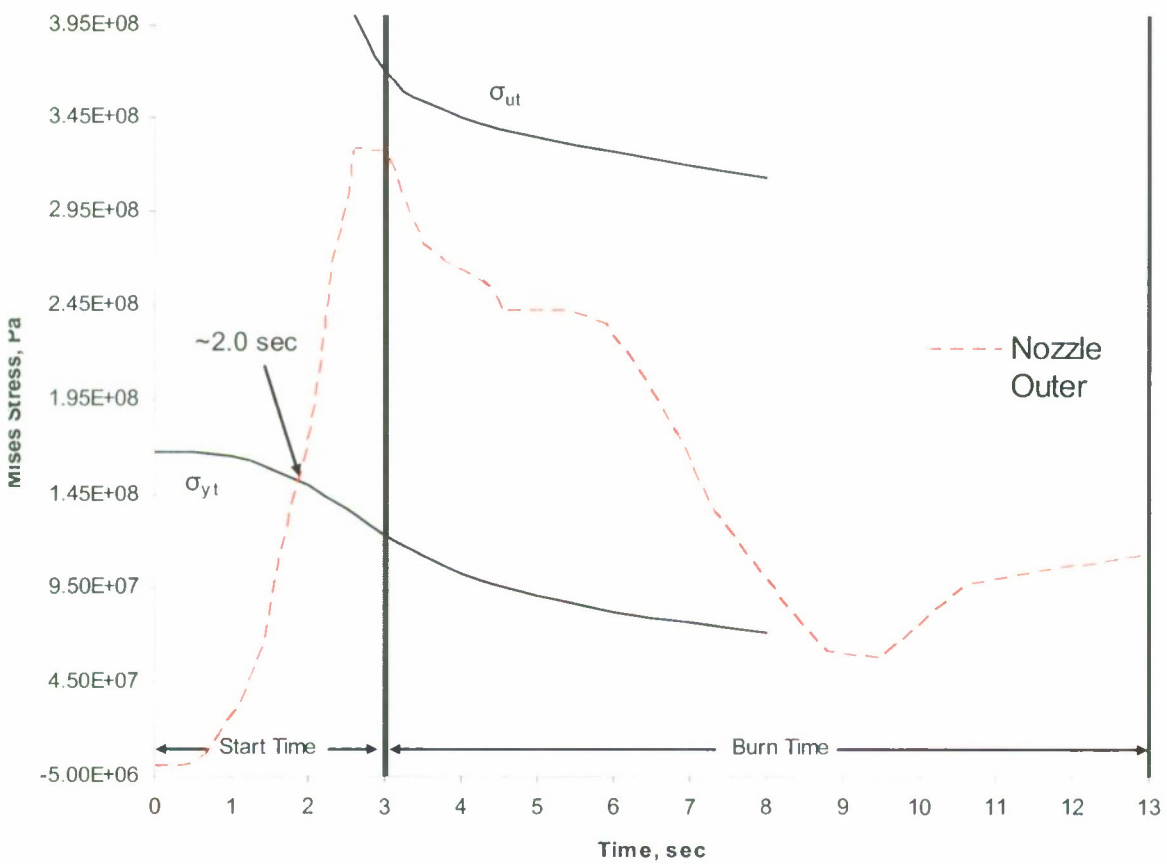


Figure 16.16 Von Mises stress versus time for 304 stainless steel employed for the nozzle outer.

The graphite nozzle liner does not seem to fair mueh better. Figure 16.17 shows an ultimate failure oeeurring before 2 seeonds. What is interesting is that the location is in a similar region of failure on conventional nozzles that have been seen in reeent tests at the Cal Poly Hybrid Roeket Motor Faeility. For the eonventional nozzles it was determined that a sharp edge at the transition from the eonical shape of the subsonic region to a cylindrieal one was resulting in a eoneentration of stress. The failure of the graphite causes a circumferential craek to propagate through the thickness of the nozzle liner. This suspieion inspired the modeling of a smooth fillet at this transition for the corresponding feature on the aerospike nozzle hardware. However, the results would seem to suggest that this still needs more attention to reduee the eoneentration of stress. A summary of results for the assembly is given in Table 16.3.

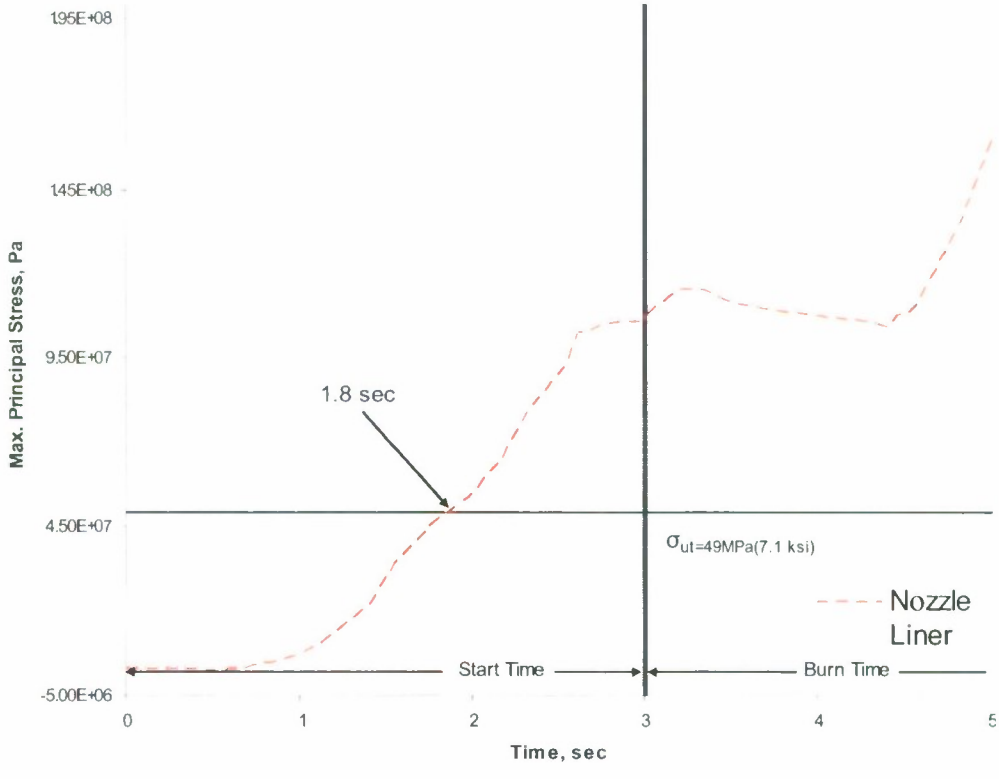


Figure 16.17 Maximum prineipal stress versus time for SiC-6 graphite employed for the nozzle liner. Plot has been truneated to 5 seeonds.

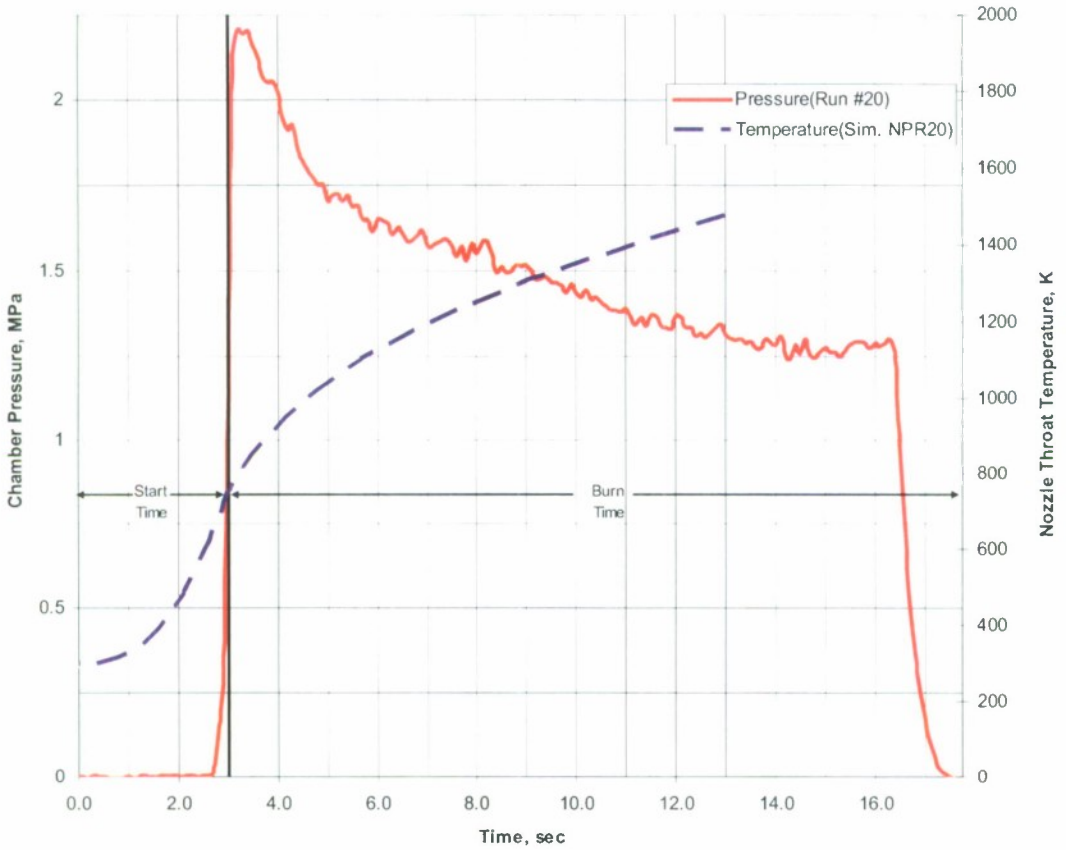


Figure 16.18 Maximum principal stress versus time for SiC-6 graphite employed for the nozzle liner. Plot has been truncated to 5 seconds.

The testing of conventional nozzles that has taken place at the Cal Poly Hybrid Rocket Motor Facility presents an opportunity to compare visual evidence of material behavior with the predictions offered here. Some sense of verification can be given to the results of the CFD and FEA models by considering how the nozzle behaves with specific regard to the nozzle liner and nozzle outer. The aforementioned circumferential cracking of the graphite liner of a conventional nozzle typically occurs during long burn times that are on the order of 10 to 15 seconds (start time not included). The FEA results presented here seem to suggest that the first instance of failure would initiate around 1.8 seconds.

At the outset it would seem that the model predicts failure prematurely, especially given the fact that the sharp edge has been eliminated at the transition from the conical

shape(defining the subsonie region) to the eylindrieal shape on the baekside of the nozzle liner. However, other groups that have sueeessfully conducted flight testing using aerospike roeket motors of the liquid and solid type have reported similar material behavior after less than 10 seeonds of total operation. In addition, to eompare the eonventional nozzle tests to the simulation of the aerospike nozzle might be flawed altogether beeause of a basie differenee in the ratio of oxidizer to fuel. Currently, in a typieal test of the eonventional roeket nozzles this ratio is approximately 8. The simulated run for the aerospike nozzle is based on a ratio of 4. It is possible for this differenee in oxidizer to fuel ratios to translate into higher Nusselt numbers for an inerease in the veloeity of the gases. Also, during testing the throat of the eonventional nozzle has been shown to erode signifieantly, inereasing the area by more than 39% for "Run#20" [21]. This makes the average chamber pressure drop and would tend to lower Nusselt number predietions based on these ehanges.

A typieal deereasing trend in pressure exhibited for these conventional nozzles is given in Figure 16.18. This data has been overlaid onto a plot of the predieted inerease in temperature with respeet to time at the nozzle liner surfaee of the throat annulus. A possible inversely proportional relationship is suggested by their eomparison. More investigation is needed to explore this relationship and there are other factors at play here that are not represented. For instanee, the simulated oxidizer-to-fuel ratio versus one that is typieally seen. Another eonsideration is that if there is more oxidizer then the effects of the oxidation eoupled with the elevated temperatures would eontribute to the erosion that effeetively lowers the ehamber pressure during the burn time.

Table 16.3 Summary of results for the thermal structural model.

Component		Material	Failure Criteria	Yield Strength,MPa (*ksi*)	Ultimate Tensile Strength, MPa(*ksi*)	Time to Yield, sec	Time to Failure, sec	Time to FS=1.25, sec
Spike Configuration #7	*Spike Tip*	*Columbium FS85*	Mises	194*(28)*	213*(31)*	2.3	2.7	2
	Spike Inner Sleeve	*Columbium FS85*	Mises	198*(29)*	251*(36)*	2.2	2.5	2
	Spike Outer Sleeve	*SiC-6 Graphite*	Max. Stress	N/A	49*(7.1)*	N/A	2.8	2.6
Spike Configuration #4	*Spike Tip*	*Columbium FS85*	Mises	187*(27)*	209*(30)*	2.3	2.8	2.1
	Spike Inner Sleeve	*Columbium FS85*	Mises	164*(24)*	209*(30)*	2.5	2.8	2.1
	Spike Outer Sleeve	*2.5D C/SiC(composite)*	Max. Stress	N/A	75.8*(11)*	N/A	2.5	1.9
Spike Configuration #1		*2.5D C/SiC(composite)*	Max. Stress	N/A	50*(7.3)*	N/A	1.7	1.6
Nozzle Liner		*SiC-6 Graphite*	Max. Stress	N/A	49*(7.1)*	N/A	1.8	1.6
Nozzle Outer		304 Stainless Steel	Mises	150*(22)*	449*(65)*	2	N/A	1.7

Spike Web Results and Discussion

The results for the spike web analysis revealed some unexpeeted but very reasonable stress coneentrations at the inner surface of the hub for the lateral loading case. When the body foree is applied to the hub of the spike web a stress is generated at the inner wall of the hub. The location of the stress coneentration is the result of the hole at the eenter of the hub being stretched by the applied body foree (Figure 16.19). The value of this stress is given in Table 16.4 along with other notable tensile stresses from this analysis. The largest failure index seems high for this stress eoncentration and represents a faetor of safety of about 1.35. In the grand seheme of things this is actually pretty good eonsidering that the loading ease was designed to be eonservative in several regards.

Table 16.4 Summary of largest prineipal stresses seen at certain loeations on the plane stress model of spike web.

Feature	*Principal Stress, MPa*	*Largest Failure Index* (σ_1/σ_{ut})
Fillet	16.4	0.565517241
Hub Inner Surface	21.5	0.74137931
Spoke	12.4	0.427586207

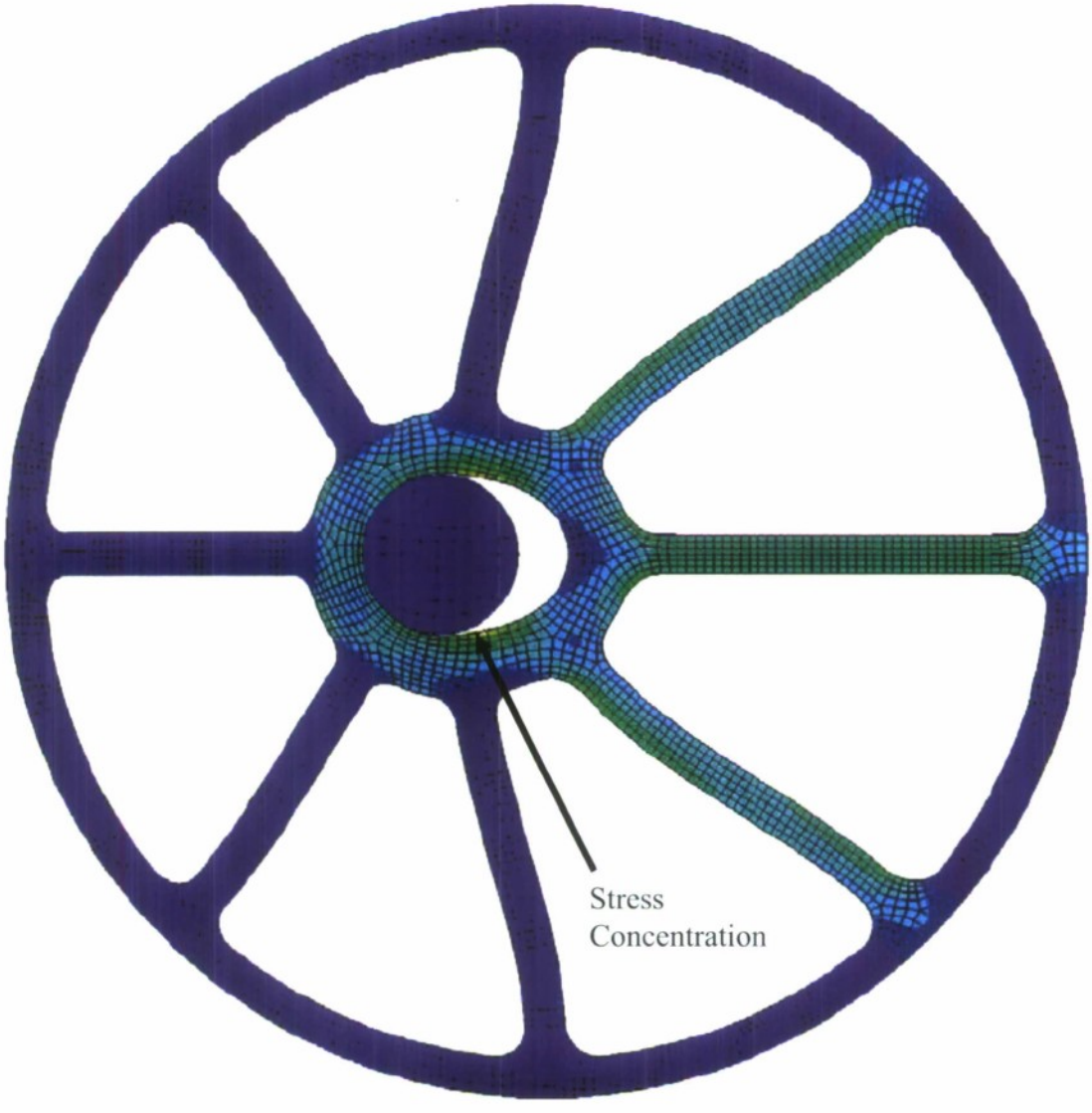


Figure 16.19 Results for one of the two loading cases for the 2-D plane stress model of the spike web.

The results of the transient heat conduction and thermal structural models are considered here separate from the lateral loading results. The imposed displacement on the spike web was applied after a free thermal expansion was allowed. The largest maximum principal stress was found to be 6.5 MPa (0.943 ksi) with the displacement boundary condition applied. This stress represents a failure index of less than 0.2 and is

considered very reasonable if the fit toleranee is held to 0.0508 mm (0.0020 in.). However, high compressive stresses were found at the fillets where the spoke transitions to the inner and outer rim. The largest eompressive stress found in these fillets was approximately 175 MPa (25.4 ksi) and represents a failure index of 1.7. The radius of these fillets will have to be redesigned to bring the failure index to within 0.5. Meeting this eriteria would require that a trade-off be made with the pereent bloeked area, as it would likely inerease when the fillets are broadened.

An alternative approaeh would be to study the response of the design to different materials. The eomparatively low conveetive coefficients at the surfaces in this region(near the inlet) result in a smaller variation of temperature with time. Using an alloy like Haynes 188 or Ineonel 625 can be considered as options for short durations where the eomponent temperature does not exeeed the design allowables. The high temperature eapabilities of these alloys would surely be the limiting faetor in the design because the stresses do not seem to appreeiable, assuming that the thermal response would be similar. Another potential drawbaek to the use of an alloy is that the rate of thermal expansion would be higher. While this may not result in failure for the spike web, it would present eomplieations for the portion of the nozzle liner that it interfaces with because of a mismatch in thermal expansion coefficients.

CHAPTER 17

Concluding Remarks and Recommendations

A multi-faceted effort in support of the design of a labseale annular aerospike nozzle has been earried out. The foeus of a preliminary design study for this effort was to assess the potential of an un-eooled eoneept of this nozzle type. In doing so, design strategies were developed with future applieations in mind. These strategies inelude the development of a CFD model to prediet gas-side surfaee film coeffieients and a method for applying these eoeffieients to an FEA model to prediet thermal stresses within the nozzle eomponents using eommereially available eodes. One notable benefit of these design tools is the ability to eompare the use of different materials. There is significant potential to advance the design effort using this work as a foundation. Several areas of this researeh can be expanded upon to help inerease the aeeuracy of future studies. Also, the results of the preliminary design study have serious implieations about the feasibility of an un-eooled nozzle design for the aerospike eonfiguration and will require that the original eoneepts, deseribed in Chapter 1, be revisited.

The analysis of an un-eooled annular aeropsike nozzle has produeed a set of predieted times at whieh the first instanee of failure oeeurs in eaeh of the nozzle eomponents. In eaeh of these eases the predieted time to failure is around 2 seconds. However, this does not mean that entire nozzle will experienee eatastrophie failure at this time. In reality, the nozzle may last mueh longer in some kind of as-damaged state. Post-flight visual inspeetions for aerospike nozzles used on solid and liquid roeket motors have revealed signifieant damage, and yet those flights yielded reasonable performance results. Testing of aetual hardware would be the next logieal step in the analysis of an

un-cooled nozzle of this type to completely understand the realistic operational limitations.

Determining the best materials to use in such hardware is essentially reduced to choosing one of the three configurations analyzed for the spike. No one configuration provided a longer predicted operating time over the others. This allows the material choice to be made based on other factors. Cost alone would suggest that configuration #7 is the most ideal. For the same cost more test specimens can be produced for configuration #7 compared to configurations #4. In addition, using an orthotropic material like 2.5D C/SiC would have non-uniform radial thermal expansion that would cause it to deviate from an axisymmetric geometry. Plus, the use of an alloy at the center of the spike would accommodate future modifications to include some form of cooling to help remove heat from the system (Table 1.1).

In the end, the biggest implication of this research is that the original aerospike design concepts need to be reexamined. With the results of this analysis and those reported for actual flight hardware in mind, the approach of an un-cooled nozzle is not adequate for a cheap, long-duration rocket motor. But the nature of a hybrid rocket motor presents an excellent opportunity to incorporate cooling. The use of nitrous-oxide as an oxidizer is a safe alternative to storing liquid oxygen for labscale testing [1]. In addition, it can also be used as a coolant in the same way liquid rockets use propllents for regenerative cooling. For the spike, this oxidizer could be bled down the center of the spike to remove heat. This would tend to waste a certain amount of the oxidizer during testing and some analysis is needed to determine what quantities are necessary to provide

effective cooling. However, if this approach is feasible, then it would be the ideal direction for the next design iteration of this nozzle type.

The design tools developed here can be directly applied to future design iterations of the annular aerospike nozzle. There are some opportunities to build upon these tools and increase the accuracy of their results. If the use of composites will be considered further, a more representative three dimensional model needs to be constructed. Doing so would allow for the orthotropic material properties to be correctly assigned in their respective orthogonal directions. Also, a more complete set of material properties for C/SiC should be cataloged if future use for this specific composite is planned. More properties should be obtained for the SiC-6 grade graphite that was studied in this research as well. The work conducted in this research is based on properties assumed to be valid for all temperatures. The FEA models developed for this study would boast a higher level of accuracy if temperature specific properties were available for these materials. The application of gap conductances should also be varied with temperature. In this research, air was assumed as the gap filling fluid for the interfacial voids created by surface roughness. Properties for air used to calculate gap conductance were taken at one elevated temperature only. In addition, the surface roughness properties for the mating components were taken as an average of typical roughnesses with regard to turning and boring methods of manufacture. Capturing more representative surface roughness and temperature dependent air properties will increase the accuracy of the gap conductance estimates. Across the variation of typical surface roughnesses alone, the gap conductance can vary by one order of magnitude. Finally, the current FEA assembly model should be changed to a 3-D sector model representing a $1/8^{th}$ slice of the actual

assembly. Such a model would be larger than the current 2-D axisymmetric model, but would allow for the inclusion of the spike web component. In this way, the assembly and spike web modeling efforts are consolidated. Other opportunities for improvement exist with regard to the CFD model that seeds the thermal loading case for the FEA model.

Perhaps the most significant improvement that can be made to the CFD model would be to make it a transient study. In this way, the gas-side surface conditions could be studied for the spike as well as the nozzle liner. The current steady-state model implies that the spike is adiabatic, therefore convective heat transfer coefficients were not predicted at the spike surface. Another benefit of modifying the CFD model to be transient is that a more accurate variation of Nusselt numbers with time can be obtained. The current research assumes that the magnitude of the axially dependent surface film coefficients varies linearly with time between the instance of ignition and 10% of the peak chamber pressure. This trend is over-predicting the actual trend during the total time of operation. An actual trend of over-prediction may even afford the un-cooled nozzle concept a second look. In addition, radiative heat transfer should be considered along with convection in the CFD model. Radiation can account for up to 35% of the total heat transfer, but the effects at the throat are not significant compared to convection [3, 7]. A closed form solution for the radiative heat transfer would be complicated and subject to large uncertainties. However, the computational benefits of a numerical model make the inclusion of the effects of radiation reasonable. The variation of gas properties was limited for the CFD and provides yet another avenue for increasing the accuracy for film coefficient results. In the model the properties of the gas were considered as "frozen" with regard to the composition of the combustion products. The only property

that was varied was the density of the gas according the ideal gas law in order to account for compressible flow. A greater accuracy with regard to local Nusselt number could be obtained if an equilibrium composition of these products was modeled using temperature specific gas properties based on the combustion analysis. Another approach would be to take advantage of Fluent's combustion analysis capabilities as an added component to the CFD model. Results for an equilibrium composition would be expected to exhibit minor variations compared that of a frozen composition, but could be appreciable if coupled with radiative heat transfer. The last topic for improving the CFD model deals with the analysis of the results. In Fluent, temperatures and heat fluxes were gathered to calculate surface film conditions based on a constant wall temperature for the aerospike nozzle. An effort was made to compare these results to semi-empirical correlations developed for a conventional nozzle on the basis of Nusselt number. A fundamental problem arose from this effort with regard to the way in which the stations were defined for the aerospike geometry. The specific problem had to do with approximating an effective hydraulic diameter for stations that did not represent planar cross sections of the nozzle. Future analysis efforts should investigate a better approach to defining these stations and the basis on which the results can compared to empirical or semi-empirical correlations.

A foundation has been laid for an effective approach to designing and analyzing an annular aerospike nozzle for labscale testing. Use has been made of commercially available CFD and FEA codes to expedite the design process. The development of design tools using these codes used an un-cooled nozzle for a preliminary design study. The results support the conclusion that an un-cooled nozzle of this type is not the best approach for long-duration burns. Furthermore, the use of C/SiC in the construction of

such a nozzle should be abandoned near the throat due to its relatively low thermal conductivity. However, this conclusion supports future research to incorporate cooling mechanisms that include, but are not limited to, a focus on removing heat transferred to the spike.

REFERENCES

1. Murray, W. and T. Carpenter, *Development of a Hybrid Rocket Motor Facility for Advanced Nozzle Research*. 2006, California State Polyteehnie University: San Luis Obispo, CA. p. 1-6.

2. Lee, C.C. and D.D. Thompson, *Fortran Program For Plug Nozzle Design*. 1964, George C. Marshall Spaee Flight Center: Hunstville, Alabama. p. 51.

3. Sutton, G.P. and O. Biblarz, eds. *Rocket Propulsion Elements*. 7th ed. 2001, John Wiley and Sons, Ine.: New York, USA.

4. Bui, T.T., et al. *Flight Research of an Aerospike Nozzle Using High Power Solid Rockets*. in *41st AIAA/ASME/SAE/ASEE Joint Propulsion Conference and Exhibit (AIAA-2005-3797)*. 2005. Tueson, Arizona.

5. Besnard, E., et al. *Design, Manufacturing and Test of a Plug Nozzle Rocket Engine*. in *38th AIAA/ASME/SAE/ASEE Joint Propulsion Conference and Exhibit(AIAA-02-4038)*. 2002. Indianapolis, Indiana.

6. Besnard, E. and J. Garvey. *Development and Flight-Testing of Liquid Propellant Aerospike Engines*. in *40th AIAA/ASME/SAE/ASEE Joint Propulsion Conference and Exhibit(AIAA-04-3354)*. 2004. Huntington Beaeh, California.

7. Shorr, M. and A.J. Zaehringer, *Solid Rocket Technology*. 1967, New York, USA: John Wiley and Sons, Ine.

8. Brown, J.A., *Materials Design Allowables Handbook For Rocket Applications*. 1981, Aerojet Liquid Roeket Company.

9. *Typical Properties of Isotropic Graphite*. 2007, Toyo Tanso Co., Ltd.

10. Shinavski, R., Various Material Properties for C/SiC and SiC Coating Speeifieations, HyperTherm High Temperature Composites, Ine., Phone Conversation, April, 2007.

11. Bunker, R.C. and A. Prince. *Hybrid Rocket Motor Nozzle Material Predictions and Results*. in *AIAA/SAE/ASME/ASEE 28th Joint Propulsion Conference and Exhibit(AIAA-1992-3591)*. 1992. Nashville, Tennessee.

12. Kowbel, W., Cost Estimates for Various High Temperature Composites, Materials and Eleetroehemieal Reseraeh Corporation, Phone Conversation, January , 2007.

13. Labruqere, et al., *Enhancement of the Oxidation resistance of the Interfacial Area in C/C composites. Part II: Oxidation Resistance of B-C, Si-B-C and Si-C Coated Carbon Preforms Densified with Carbon.* Journal of the European Ceramie Soeiety, 2002. **22**: p. 1011-1021.

14. Jones, R.M., ed. *Mechanics of Composite Materials*. 2nd ed. 1999, Brunner-Routledge: New York, USA.

15. Dalmaz, A., et al., *Elastic Moduli of a 2.5D C/SiC Composite: Experimental and Theoretical Estimates.* Composites Seienee and Teehnology, 2000. **60**: p. 913-925.

16. *Chemical Vapor Infiltration Products*. 2006, General Electrie Infrastrueture.

17. *Fluent Version 6.2.16 User's Guide, 2005.*

18. Thorneroft, G., Leetures on Fully Developed Turbulent Flow, ME553 Convection Heat Transfer, Winter 2007.

19. Shollenberger, K., Leetures on CFD Modeling, ME554 Computational Heat Transfer, Winter 2006.

20. Spivey, S., *Geometric Effects On Pressure Drop*, in *Mechanical Engineering*. 2006, California State Polyteehnie University: San Luis Obispo, CA. p. 14-16.

21. Carpenter, T., Leetures on Roeket Propulsion, ME541 Advaneed Thermodynamies, Fall 2006.

22. Ineropera, F.P. and D.P. DeWitt, eds. *Fundamentals of Heat and Mass Transfer*. 4th ed. 1996, John Wiley and Sons, Inc.: New York, NY.

23. Barrere, M., et al., *Rocket Propulsion*. 1960, Amsterdam, Netherlands: Elsevier Publishing Company.

24. Bartz, D.R., *A Simple Equation for Rapid Estimation of Rocket Nozzle Convective Heat Transfer Coefficients*. Jet Propulsion, 1957. **27**(1): p. 49-51.

25. Fox, R.W., A.T. McDonald, and P.J. Pritchard, eds. *Introduction to Fluid Mechanics*. 6th ed. 2006, John Wiley and Sons, Inc.: New York, NY.

26. Sehetz, J.A., *Boundary Layer Analysis*. 1993, New Jersey, USA: Prentiee Hall.

27. Ahmad, R.A., *Convective Heat Transfer in the Reusable Solid Rocket Motor of the Space Transportation System.* Heat Transfer Engineering, 2005. **26**(10): p. 30-45.

28. Mello, J., Per Working Papers, 2007: San Luis Obispo, CA.

29. *Abaqus Version 6.6 Documentation, 2007.*

30. Chung, S.Y. and H.J. Sung, *Direct Numerical Simulation of Turbulent Concentric Annular Pipe Flow, Part 2: Heat Transfer.* International Journal of Heat and Fluid Flow, 2003. **24**: p. 399-411.

31. Savija, I., et al. *Review of Thermal Conductance Models For Joints Incorporating Enhancement Materials.* in *40th Annual Aerospce Sciences Meeting And Exhibit (AIAA-2002-0494).* 2002. Reno, Nevada.

32. Hossain, M.R., *Optimization of Heat Sinks with Flow Bypass Using Entropy Generation Minimization*, in *Mechanical Engineering.* 2006, University of Waterloo. p. 207.

33. Holman, J.P., ed. *Heat Transfer.* 5th ed. 1981, McGraw-Hill Book Company: New York, USA.

34. Oberg, E., et al., eds. *Machinery's Handbook.* 26th ed. 2000, Industrial Press, Inc.: New York, USA.

35. Cook, R.D. and W.C. Young, eds. *Advanced Mechanics of Materials.* 2nd ed. 1999, Prentice Hall: Upper Saddle River, NJ.

36. Timoshenko, S. and J.N. Goodier, eds. *Theory of Elasticity.* 2nd ed. 1951, McGraw-Hill Book Company: New York, USA.

APPENDIX A

Run#20 Data

The following is basic data taken for a test firing of a conventional converging-diverging nozzle. This data supported the lectures of Dr. Tom Carpenter at California State Polytechnic University in San Luis Obispo for ME541 Advanced Thermodynamics/Rocket Propulsion.

----------Run #20----------

Data from the actual test firing

The average chamber pressure and burn time were determined using the filtered data for the run.

$A_t := .2911\,in^2$ Average throat area

$P_c := 217.1882\,psi$ Average chamber pressure $t_b := 13.556\,s$ Burn time

The mass flow rate of the oxidizer was found by weighing the Nitrous bottle before and after the test.

Oxidizer bottle mass (initial) $m_{oxbottle0} := 196.8\,lb$

Oxidizer bottle mass (final) $m_{oxbottlef} := 191.0\,lb$

$\Delta m_{ox} := m_{oxbottle0} - m_{oxbottlef}$

Mass flow rate of the oxidizer:

$m_{dotox} := \frac{\Delta m_{ox}}{t_b}$ $m_{dotox} = 0.428\,\frac{1}{s}\,lb$ $m_{dotox} = 0.194\,\frac{1}{s}\,kg$

Fuel grain mass (initial) $m_{fuel0} := 547.5\,gm$

Fuel grain mass (final) $m_{fuelf} := 252.75\,gm$

$\Delta m_{fuel} := m_{fuel0} - m_{fuelf}$

$\Delta m_{fuel} = 0.65\,lb$

Mass flow rate of the fuel:

$m_{dotfuel} := \frac{\Delta m_{fuel}}{t_b}$ $m_{dotfuel} = 0.0479\,\frac{1}{s}\,lb$

$m_{dotfuel} = 0.0217\,\frac{1}{s}\,kg$

Mixture Ratio(oxidizer/fuel)

$MR := \frac{m_{dotox}}{m_{dotfuel}}$ $MR = 8.92565$

From TEP, based on an average exit area ratio of 1.8718 for the average throat area

Stagnation Temperature	Characteristic Velocity	Thrust Coefficient	Specific Impulse
$T_0 := 5374\,R$	$C_{star} := 4757\,\frac{ft}{s}$	$C_f := 1.190$	$I_{sp} := 175.88\,\frac{lbf \cdot s}{lb}$

NOTE: This data came from a combustion analysis executed in TEP(Thermal Equilibrium Program) and is based on an average area ratio at the throat. The throat area increased by 39.5% during the test affecting performance.

APPENDIX B

Base Case Study Gas Properties

Source File:

Base Case Gas Properties.xls

Notes:

1. Includes gas properties generated using *TEP* and Nusselt number calculations for each station in the base case study.
2. Contains results from CFD Study.

APPENDIX C

Aerospike Study Gas Properties

Source File:

Aerospike Gas Properties.xls

Notes:

1. Includes gas properties generated using *TEP*, Nusselt number and convective heat transfer coefficient calculations for each station in the aerospike study.
2. Contains results from CFD study.

APPENDIX D

CFD Study: Supporting Fluent Files

Source Files:

- ***Grid Density Study:***
  - *Base_nozzle_1890_gds.cas*
  - *Base_nozzle_1890_gds.dat*
  - *Base_nozzle_4590_gds.cas*
  - *Base_nozzle_4590_gds.dat*
  - *Base_nozzle_6390_gds.cas*
  - *Base_nozzle_6390_gds.dat*
  - *Base_nozzle_8190_gds.cas*
  - *Base_nozzle_8190_gds.dat*
- ***Base Case:***
  - *Base_nozzle_8190.cas*
  - *Base_nozzle_8190.dat*
- ***Aerospike***
  - *AS_nozzle_8190_gds.cas*
  - *ASe_nozzle_8190_gds.dat*

Notes:

- *Base=Base Case Study*
- *AS=Aerospike Study*
- *gds=grid density*
- *1890, 4590, 6390, 8190=Grid Density*

APPENDIX E

Abaqus FILM User-Subroutine

Source Files:

- ***Assembly Model***
  - *NPR_20_Gas-Side2.for*
- ***Spike Web Model***
  - *NPR_20_Gas-Side3.for*

Notes: *When applying these user-subroutines make sure that the selected surfaces match the selected surfaces in the input file for the job.*

APPENDIX F

FEA Verification Study: Sample Calculations

Simple Steady State 1-D heat transfer solution: Assumes a cylinder with graphite inner sleeve and stainless steel outer sleeve. This model will be used to verfiy results for an FEA model in Abaqus CAE.

$h_{throat} := 7000\frac{W}{m^2 K}$ coefficient of heat transfer for the combusted gas

$k_{SiC6} := 130\frac{W}{m \cdot K}$ thermal conductivity for SiC-6 graphite

$k_{SiC12} := 80\frac{W}{m \cdot K}$ thermal conductivity for SiC-12 graphite

$k_{SS304} := 21.96\frac{W}{m \cdot K}$ thermal conductivity for 304 Stainless Steel(@ 812 K)

$h_{natural} := 5\frac{W}{m^2 K}$ coefficient of heat transfer for the combusted gas

$T_{aw} := 3000K$ adiabatic wall temperature of the fluid at the throat

$T_{amb} := 300K$ assumed ambient temperature

$r_1 := .02m$ radius to the graphite wall

$r_2 := .03m$ radius to the interaction between the graphite and stainless steel

$r_3 := .04m$ radius to the outside of the 304 Stainless Steel

$L := .044m$ Length of the compound cylinder

Assuming no thermal barrier in between the stainless steel and the graphite

$$q_{rate} := \left[\frac{1}{h_{throat} \cdot 2 \cdot \pi r_1 \cdot L} + \frac{\ln\left(\frac{r_2}{r_1}\right)}{k_{SiC6}(2 \cdot \pi \cdot L)} + \frac{\ln\left(\frac{r_3}{r_2}\right)}{k_{SS304}(2 \cdot \pi \cdot L)} + \frac{1}{h_{natural}(2 \cdot \pi r_3 \cdot L)}\right]^{-1} \cdot (T_{aw} - T_{amb})$$

$q_{rate} = 148.594W$

Heat fluxes across the Graphite inner surface and the SS outer surface

$$q_{flux1} := \frac{q_{rate}}{2 \cdot \pi \cdot r_1 \cdot L}$$ Flux across the graphite surface

$$q_{flux1} = 2.687 \times 10^4 \frac{W}{m^2}$$

$$q_{flux2} := \frac{q_{rate}}{2 \cdot \pi \cdot r_3 \cdot L}$$ Flux across the SS surface

$$q_{flux2} = 1.344 \times 10^4 \frac{W}{m^2}$$

Temperatures at the gas-side wall, interface and the outer wall.

$$T_1 := T_{aw} - \frac{q_{rate}}{h_{throat} \cdot 2 \cdot \pi \cdot r_1 \cdot L}$$ Inner wall temperature

$$T_1 = 2.996 \times 10^3 \text{ K}$$

$$T_2 := T_1 - q_{rate} \cdot \frac{\ln\left(\frac{r_2}{r_1}\right)}{k_{SiC6}(2 \cdot \pi \cdot L)}$$ Interface temperature

$$T_2 = 2.994 \times 10^3 \text{ K}$$

$$T_3 := T_2 - q_{rate} \cdot \frac{\ln\left(\frac{r_3}{r_2}\right)}{k_{SS304}(2 \cdot \pi \cdot L)}$$ Outer wall temperature

$$T_3 = 2.987 \times 10^3 \text{ K}$$

Simple Transient 1-D heat transfer solution: Assumes a hollow cylinder made of stainless steel with an insulating graphite liner. The calculations are based on an example from "Rocket Propulsion"[23]. In this example the hollow cylinder represent a combustion chamber. This model will be used to verfiy results for an FEA model in Abaqus CAE.

The cylindrical wall is assumed to be approximately a flat plate because the cylinder thickness is small relative to chamber diameter.

Material Properties

$T_i := 300K$ Initial temperature

$k_{SiC6} = 130\frac{m \cdot kg}{K \cdot s^3}$ $C_{pSiC6} := 710\frac{J}{kg \cdot K}$ $L_c := \frac{.01}{2}m$

$\rho_{SiC6} := 1850\frac{kg}{m^3}$ $\alpha_{SiC6} := \frac{k_{SiC6}}{\rho_{SiC6} C_{pSiC6}}$

Fluid Properties

$T_{inf} := 3000K$ Temperature of the fluid

$h_{arb} := 5200\frac{W}{m^2 \cdot K}$

Burn time

$t_b := 1sec, 2sec .. 30sec$

Biot and Fourier Numbers

$Bi := \frac{h_{arb} \cdot L_c}{k_{SiC6}}$ $Bi = 0.2$ $F_0(t_b) := \alpha_{SiC6}\frac{t_b}{L_c^2}$

$C_1 := 1.0311$ $\zeta_1 := .4328rad$

Dimensionless temp calculations for the outer the wall.(normally this would be to the center of the thickness of a plate, but the thin wall assumption assumes a uniform temperature distribution)

$\theta_0(t_b) := C_1 \cdot e^{-\zeta_1^2 \cdot F_0(t_b)}$

Temperature at the outer wall after burn time.

$T_0(t_b) := \theta_0(t_b) \cdot (T_i - T_{inf}) + T_{inf}$

Thermal Stress Analysis: Long Circular cylinder(Plane Strain) from "Theory of Elasticity"[36]

$b := r_3 \quad a := r_2 \quad \nu := .29 \quad E := 200000000000\frac{N}{m^2} \quad \alpha := 18.7\cdot10^{-6}\frac{m}{m\cdot K}$

@ r=a $\quad r := a$

$T_i := T_2 - T_3 \qquad T := \frac{T_i}{\log\left(\frac{b}{a}\right)}\cdot\log\left(\frac{b}{r}\right)$

$$\sigma_r := \frac{\alpha\cdot E\cdot T_i}{2\cdot(1-\nu)\cdot\log\left(\frac{b}{a}\right)}\cdot\left[-\log\left(\frac{b}{r}\right) - \frac{a^2}{b^2-a^2}\cdot\left(1-\frac{b^2}{r^2}\right)\cdot\log\left(\frac{b}{a}\right)\right]$$

$$\sigma_\theta := \frac{\alpha\cdot E\cdot T_i}{2\cdot(1-\nu)\cdot\ln\left(\frac{b}{a}\right)}\cdot\left(1-\frac{2\cdot b^2}{b^2-a^2}\cdot\ln\left(\frac{b}{a}\right)\right)$$

$$\sigma_z := \frac{\alpha\cdot E\cdot T_i}{2\cdot(1-\nu)\cdot\ln\left(\frac{b}{a}\right)}\cdot\left(1-\frac{2\cdot b^2}{b^2-a^2}\cdot\ln\left(\frac{b}{a}\right)\right)$$

$\sigma_r = 0\,Pa$ $\qquad \sigma_\theta = -2.031\times10^7\,Pa$ $\qquad \sigma_z = -2.031\times10^7\,Pa$

@ r=b $\quad r := b$

$T_i := T_2 - T_3 \qquad T := \frac{T_i}{\log\left(\frac{b}{a}\right)}\cdot\log\left(\frac{b}{r}\right)$

$$\sigma_r := \frac{\alpha\cdot E\cdot T_i}{2\cdot(1-\nu)\cdot\log\left(\frac{b}{a}\right)}\cdot\left[-\log\left(\frac{b}{r}\right) - \frac{a^2}{b^2-a^2}\cdot\left(1-\frac{b^2}{r^2}\right)\cdot\log\left(\frac{b}{a}\right)\right]$$

$$\sigma_\theta := \frac{\alpha\cdot E\cdot T_i}{2\cdot(1-\nu)\cdot\ln\left(\frac{b}{a}\right)}\cdot\left(1-\frac{2\cdot a^2}{b^2-a^2}\cdot\ln\left(\frac{b}{a}\right)\right)$$

$$\sigma_z := \frac{\alpha\cdot E\cdot T_i}{2\cdot(1-\nu)\cdot\ln\left(\frac{b}{a}\right)}\cdot\left(1-\frac{2\cdot a^2}{b^2-a^2}\cdot\ln\left(\frac{b}{a}\right)\right)$$

$\sigma_r = 0\,Pa$ $\qquad \sigma_\theta = 1.678\times10^7\,Pa$ $\qquad \sigma_z = 1.678\times10^7\,Pa$

@ r=(b+a)/2 $r := \frac{b + a}{2}$

$T_i := T_2 - T_3$ $T := \frac{T_i}{\log\left(\frac{b}{a}\right)} \cdot \log\left(\frac{b}{r}\right)$

$$\sigma_r := \frac{\alpha \cdot E \cdot T_i}{2 \cdot (1 - \nu) \cdot \ln\left(\frac{b}{a}\right)} \cdot \left[-\ln\left(\frac{b}{r}\right) - \frac{a^2}{b^2 - a^2} \cdot \left(1 - \frac{b^2}{r^2}\right) \cdot \ln\left(\frac{b}{a}\right) \right]$$

$$\sigma_\theta := \frac{\alpha \cdot E \cdot T_i}{2 \cdot (1 - \nu) \cdot \ln\left(\frac{b}{a}\right)} \cdot \left[1 - \ln\left(\frac{b}{r}\right) - \frac{a^2}{b^2 - a^2} \cdot \left(1 + \frac{b^2}{r^2}\right) \cdot \ln\left(\frac{b}{a}\right) \right]$$

$$\sigma_z := \frac{\alpha \cdot E \cdot T_i}{2 \cdot (1 - \nu) \cdot \ln\left(\frac{b}{a}\right)} \cdot \left(1 - 2\ln\left(\frac{b}{r}\right) - \frac{2 \cdot a^2}{b^2 - a^2} \cdot \ln\left(\frac{b}{a}\right) \right)$$

$\sigma_r = -1.309 \times 10^6$ Pa $\sigma_\theta = 8.694 \times 10^5$ Pa $\sigma_z = -4.394 \times 10^5$ Pa

Pressurized Cylinder Stress Analysis: Circular cylinder(Plane Stress)

$p_i := 300\text{psi}$ $p_o := 14.7\text{psi}$

$b := r_2$ $a := r_3$

@ $r := a$

$$\sigma_r := p_i \cdot \frac{b^2}{a^2 - b^2} \cdot \left(1 - \frac{a^2}{r^2}\right) - p_o \cdot \left[\frac{a^2}{a^2 - b^2} \cdot \left(1 - \frac{b^2}{r^2}\right)\right]$$

$$\sigma_\theta := p_i \cdot \frac{b^2}{a^2 - b^2} \cdot \left(1 + \frac{a^2}{r^2}\right) - p_o \cdot \left[\frac{a^2}{a^2 - b^2} \cdot \left(1 + \frac{b^2}{r^2}\right)\right]$$

$\sigma_r = -1.014 \times 10^5\ \text{Pa}$ $\sigma_\theta = 4.957 \times 10^6\ \text{Pa}$

@ $r := b$

$$\sigma_r := p_i \cdot \frac{b^2}{a^2 - b^2} \cdot \left(1 - \frac{a^2}{r^2}\right) - p_o \cdot \left[\frac{a^2}{a^2 - b^2} \cdot \left(1 - \frac{b^2}{r^2}\right)\right]$$

$$\sigma_\theta := p_i \cdot \frac{b^2}{a^2 - b^2} \cdot \left(1 + \frac{a^2}{r^2}\right) - p_o \cdot \left[\frac{a^2}{a^2 - b^2} \cdot \left(1 + \frac{b^2}{r^2}\right)\right]$$

$\sigma_r = -2.068 \times 10^6\ \text{Pa}$ $\sigma_\theta = 6.924 \times 10^6\ \text{Pa}$

APPENDIX G

C/SiC Material Stiffness Matrix

(all stiffnesses in units of MPa)

----------High Temperature----------

Estimate of material stiffnesses for ideally non-cracked 2.5D C/SiC composite(from table 8 of "elastic moduli of a 2.5 D C/SiC composite: experimental and theoretical estimates")[15], NOTE: ideally noncracked would suggest high temperature state. The assumed high temperature is 2273K.

$f := .85$ knock down factor for wavvyness

Material stiffness matrix, no material rotation

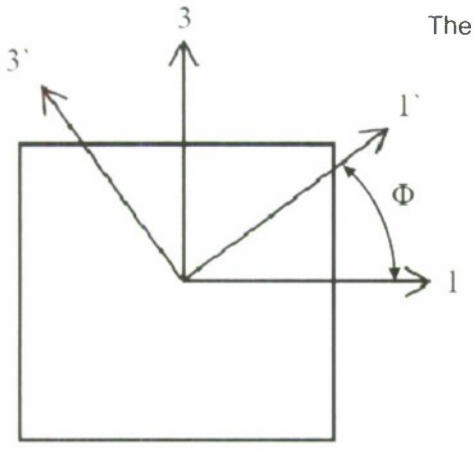
$$C := \begin{pmatrix} f\cdot 105 & 15.1 & 11.1 & 0 & 0 & 0 \\ 15.1 & f\cdot 105 & 11.1 & 0 & 0 & 0 \\ 11.1 & 11.1 & 36.6 & 0 & 0 & 0 \\ 0 & 0 & 0 & 13.5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 13.5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 24.1 \end{pmatrix} \quad C = \begin{pmatrix} 89.25 & 15.1 & 11.1 & 0 & 0 & 0 \\ 15.1 & 89.25 & 11.1 & 0 & 0 & 0 \\ 11.1 & 11.1 & 36.6 & 0 & 0 & 0 \\ 0 & 0 & 0 & 13.5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 13.5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 24.1 \end{pmatrix}$$

$\phi := 90\text{deg}$

$$T(\phi) := \begin{pmatrix} \cos(\phi)^2 & 0 & \sin(\phi)^2 & 0 & 0 & \sin(\phi)\cos(\phi) \\ 0 & 1 & 0 & 0 & 0 & 0 \\ \sin(\phi)^2 & 0 & \cos(\phi)^2 & 0 & 0 & -\sin(\phi)\cos(\phi) \\ 0 & 0 & 0 & \cos(\phi) & 0 & -\sin(\phi) \\ 0 & 0 & 0 & 0 & \cos(\phi)^2 - \sin(\phi)^2 & 0 \\ -2\sin(\phi)\cos(\phi) & 0 & 2\sin(\phi)\cos(\phi) & \sin(\phi) & 0 & \cos(\phi) \end{pmatrix}$$

$C' := T(\phi)^T \, C \cdot T(\phi)$ For a 90° material rotation about the 2-axis(out of the page)

$$C' = \begin{pmatrix} 36.6 & 11.1 & 11.1 & -2.951\times 10^{-15} & 0 & -1.561\times 10^{-15} \\ 11.1 & 89.25 & 15.1 & 0 & 0 & 0 \\ 11.1 & 15.1 & 89.25 & 2.951\times 10^{-15} & 0 & 4.785\times 10^{-15} \\ -2.951\times 10^{-15} & 0 & 2.951\times 10^{-15} & 24.1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 13.5 & 0 \\ -1.561\times 10^{-15} & 0 & 4.785\times 10^{-15} & 0 & 0 & 13.5 \end{pmatrix}$$



The prime superscript denotes rotated configuration.

Solving for the poisson's ratios

1)First we need the compliance matrix

$$\begin{pmatrix} S_{11} & S_{12} & S_{13} & S_{14} & S_{15} & S_{16} \\ S_{21} & S_{22} & S_{23} & S_{24} & S_{25} & S_{26} \\ S_{31} & S_{32} & S_{33} & S_{34} & S_{35} & S_{36} \\ S_{41} & S_{42} & S_{43} & S_{44} & S_{45} & S_{46} \\ S_{51} & S_{52} & S_{53} & S_{54} & S_{55} & S_{56} \\ S_{61} & S_{62} & S_{63} & S_{64} & S_{65} & S_{66} \end{pmatrix} := C^{-1}$$

Corresponding Compliance Matrix

2) Now we need initial guesses for the poisson's ratios

$$\nu_{21} := .03 \quad \nu_{31} := .03 \quad \nu_{23} := .03 \quad \nu_{12} := .0116 \quad \nu_{13} := .03 \quad \nu_{32} := .03$$

3) The setup

Equation 2.35 from Jones, Mechanics of Composite Materials[14]

$$S := S_{11} \cdot S_{22} \cdot S_{33} - S_{11} \cdot S_{23}^{2} - S_{22} \cdot S_{13}^{2} - S_{33} \cdot S_{12}^{2} + 2 \cdot S_{12} \cdot S_{23} \cdot S_{13}$$

$$S = 3.775 \times 10^{-6}$$

Relations for Elastic Modulus

$$E_1 := \frac{1}{S_{11}} \qquad E_2 := E_1 \qquad E_3 := \frac{1}{S_{33}}$$

4) The "given" statement is required for the solver, note the equations have the boolean equals sign. The six expressions that follow come directly from the equation 2.36 in Jones, Mechanics of Composite Materials.

Given

$$f(105) = \frac{1 - \nu_{23} \cdot \nu_{32}}{E_2 \cdot E_3 \cdot \dfrac{1 - \nu_{12} \cdot \nu_{21} - \nu_{23} \cdot \nu_{32} - \nu_{31} \cdot \nu_{13} - 2 \cdot \nu_{21} \cdot \nu_{32} \cdot \nu_{13}}{E_1 \cdot E_2 \cdot E_3}}$$

$$f(105) = \frac{1 - \nu_{13} \cdot \nu_{31}}{E_1 \cdot E_3 \cdot \dfrac{1 - \nu_{12} \cdot \nu_{21} - \nu_{23} \cdot \nu_{32} - \nu_{31} \cdot \nu_{13} - 2 \cdot \nu_{21} \cdot \nu_{32} \cdot \nu_{13}}{E_1 \cdot E_2 \cdot E_3}}$$

$$15.1 = \frac{\nu_{21} + \nu_{31}\cdot\nu_{23}}{E_2\cdot E_3\cdot\dfrac{1-\nu_{12}\cdot\nu_{21}-\nu_{23}\cdot\nu_{32}-\nu_{31}\cdot\nu_{13}-2\cdot\nu_{21}\cdot\nu_{32}\cdot\nu_{13}}{E_1\cdot E_2\cdot E_3}}$$

$$11.1 = \frac{\nu_{32} + \nu_{12}\cdot\nu_{31}}{E_1\cdot E_3\cdot\dfrac{1-\nu_{12}\cdot\nu_{21}-\nu_{23}\cdot\nu_{32}-\nu_{31}\cdot\nu_{13}-2\cdot\nu_{21}\cdot\nu_{32}\cdot\nu_{13}}{E_1\cdot E_2\cdot E_3}}$$

$$11.1 = \frac{\nu_{31} + \nu_{21}\cdot\nu_{32}}{E_2\cdot E_3\cdot\dfrac{1-\nu_{12}\cdot\nu_{21}-\nu_{23}\cdot\nu_{32}-\nu_{31}\cdot\nu_{13}-2\cdot\nu_{21}\cdot\nu_{32}\cdot\nu_{13}}{E_1\cdot E_2\cdot E_3}}$$

$$36.6 = \frac{1-\nu_{12}\cdot\nu_{21}}{E_1\cdot E_2\cdot\dfrac{1-\nu_{12}\cdot\nu_{21}-\nu_{23}\cdot\nu_{32}-\nu_{31}\cdot\nu_{13}-2\cdot\nu_{21}\cdot\nu_{32}\cdot\nu_{13}}{E_1\cdot E_2\cdot E_3}}$$

$$\begin{pmatrix}\nu_{21val}\\ \nu_{31val}\\ \nu_{23val}\\ \nu_{12val}\\ \nu_{13val}\\ \nu_{32val}\end{pmatrix} := \text{Find}\left(\nu_{21},\nu_{31},\nu_{23},\nu_{12},\nu_{13},\nu_{32}\right) \qquad \begin{pmatrix}\nu_{21val}\\ \nu_{31val}\\ \nu_{23val}\\ \nu_{12val}\\ \nu_{13val}\\ \nu_{32val}\end{pmatrix} = \begin{pmatrix}0.137\\ 0.106\\ 0.262\\ 0.137\\ 0.262\\ 0.106\end{pmatrix}$$

Reciprocal relations check

$$\frac{\nu_{21val}}{E_2} = 1.621\times 10^{-3}$$

$$\frac{\nu_{12val}}{E_1} = 1.621\times 10^{-3}$$

$$\frac{\nu_{31val}}{E_3} = 3.107\times 10^{-3} \qquad \frac{\nu_{32val}}{E_3} = 3.107\times 10^{-3}$$

$$\frac{\nu_{13val}}{E_1} = 3.107\times 10^{-3} \qquad \frac{\nu_{23val}}{E_2} = 3.107\times 10^{-3}$$

Resulting Engineering Constants @ High Temperature

$E_1 = 84.281$

$E_2 = 84.281$

$E_3 = 34.239$

$G_{23} := \frac{1}{S_{44}}$ $G_{23} = 13.5$

$G_{31} := \frac{1}{S_{55}}$ $G_{31} = 13.5$

$G_{12} := \frac{1}{S_{66}}$ $G_{12} = 24.1$

$\nu_{21} := \nu_{21val}$ $\nu_{21} = 0.137$

$\nu_{31} := \nu_{31val}$ $\nu_{31} = 0.106$

$\nu_{23} := \nu_{23val}$ $\nu_{23} = 0.262$

----------Low Temperature----------

Estimate of the material stiffnesses for 2.5D C/SiC composite at room temperature with assumed flattened tows(from table 7a "elastic moduli of a 2.5 D C/SiC composite: experimental and theoretical estimates")[15], NOTE: these properties will be knocked down by 15%, as suggested in the paper, to account for the "wavvyness" of the tows.

$f := .85$ knock down factor for wavvyness

Stiffness matrix, no material rotation

$$C := \begin{pmatrix} f\cdot 89 & 8.7 & 7.7 & 0 & 0 & 0 \\ 8.7 & f\cdot 89 & 7.7 & 0 & 0 & 0 \\ 7.7 & 7.7 & 31.5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 11.3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 11.3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 19.6 \end{pmatrix} \qquad C = \begin{pmatrix} 75.65 & 8.7 & 7.7 & 0 & 0 & 0 \\ 8.7 & 75.65 & 7.7 & 0 & 0 & 0 \\ 7.7 & 7.7 & 31.5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 11.3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 11.3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 19.6 \end{pmatrix}$$

$\phi := 90\text{deg}$

$$T(\phi) := \begin{pmatrix} \cos(\phi)^2 & 0 & \sin(\phi)^2 & 0 & 0 & \sin(\phi)\cos(\phi) \\ 0 & 1 & 0 & 0 & 0 & 0 \\ \sin(\phi)^2 & 0 & \cos(\phi)^2 & 0 & 0 & -\sin(\phi)\cos(\phi) \\ 0 & 0 & 0 & \cos(\phi) & 0 & -\sin(\phi) \\ 0 & 0 & 0 & 0 & \cos(\phi)^2 - \sin(\phi)^2 & 0 \\ -2\sin(\phi)\cos(\phi) & 0 & 2\sin(\phi)\cos(\phi) & \sin(\phi) & 0 & \cos(\phi) \end{pmatrix}$$

$C' := T(\phi)^T \cdot C \cdot T(\phi)$ For a 90° material rotation about the 2-axis

$$C' = \begin{pmatrix} 31.5 & 7.7 & 7.7 & -2.4\times 10^{-15} & 0 & -1.457\times 10^{-15} \\ 7.7 & 75.65 & 8.7 & 0 & 0 & 0 \\ 7.7 & 8.7 & 75.65 & 2.4\times 10^{-15} & 0 & 4.161\times 10^{-15} \\ -2.4\times 10^{-15} & 0 & 2.4\times 10^{-15} & 19.6 & 0 & 0 \\ 0 & 0 & 0 & 0 & 11.3 & 0 \\ -1.457\times 10^{-15} & 0 & 4.161\times 10^{-15} & 0 & 0 & 11.3 \end{pmatrix}$$

Solving for the poisson's ratios

1)First we need the compliance matrix

$$\begin{pmatrix} S_{11} & S_{12} & S_{13} & S_{14} & S_{15} & S_{16} \\ S_{21} & S_{22} & S_{23} & S_{24} & S_{25} & S_{26} \\ S_{31} & S_{32} & S_{33} & S_{34} & S_{35} & S_{36} \\ S_{41} & S_{42} & S_{43} & S_{44} & S_{45} & S_{46} \\ S_{51} & S_{52} & S_{53} & S_{54} & S_{55} & S_{56} \\ S_{61} & S_{62} & S_{63} & S_{64} & S_{65} & S_{66} \end{pmatrix} := C^{-1}$$

2) Now we need initial guesses for the poisson's ratios

$\nu_{21} := .03 \quad \nu_{31} := .03 \quad \nu_{23} := .03 \quad \nu_{12} := .0116 \quad \nu_{13} := .03 \quad \nu_{32} := .03$

3) The setup

Equation 2.35 from Jones, Mechanics of Composite Materials[14]

$$S := S_{11} \cdot S_{22} \cdot S_{33} - S_{11} \cdot S_{23}^{2} - S_{22} \cdot S_{13}^{2} - S_{33} \cdot S_{12}^{2} + 2 \cdot S_{12} \cdot S_{23} \cdot S_{13}$$

$$S = 5.884 \times 10^{-6}$$

Relations for Elastic Modulus

$$E_1 := \frac{1}{S_{11}} \qquad E_2 := E_1 \qquad E_3 := \frac{1}{S_{33}}$$

4) The "given" statement is required for the solver, not the equations have the boolean equals sign. The six expressions that follow come directly from the equation 2.36 in Jones, Mechanics of Composite Materials.

Given

$$f \cdot 89 = \frac{1 - \nu_{23} \cdot \nu_{32}}{E_2 \cdot E_3 \cdot \frac{1 - \nu_{12} \cdot \nu_{21} - \nu_{23} \cdot \nu_{32} - \nu_{31} \cdot \nu_{13} - 2 \cdot \nu_{21} \cdot \nu_{32} \cdot \nu_{13}}{E_1 \cdot E_2 \cdot E_3}}$$

$$f \cdot 89 = \frac{1 - \nu_{13} \cdot \nu_{31}}{E_1 \cdot E_3 \cdot \frac{1 - \nu_{12} \cdot \nu_{21} - \nu_{23} \cdot \nu_{32} - \nu_{31} \cdot \nu_{13} - 2 \cdot \nu_{21} \cdot \nu_{32} \cdot \nu_{13}}{E_1 \cdot E_2 \cdot E_3}}$$

$$8.7 = \frac{\nu_{21} + \nu_{31}\cdot\nu_{23}}{E_2\cdot E_3\cdot\dfrac{1 - \nu_{12}\cdot\nu_{21} - \nu_{23}\cdot\nu_{32} - \nu_{31}\cdot\nu_{13} - 2\cdot\nu_{21}\cdot\nu_{32}\cdot\nu_{13}}{E_1\cdot E_2\cdot E_3}}$$

$$7.7 = \frac{\nu_{32} + \nu_{12}\cdot\nu_{31}}{E_1\cdot E_3\cdot\dfrac{1 - \nu_{12}\cdot\nu_{21} - \nu_{23}\cdot\nu_{32} - \nu_{31}\cdot\nu_{13} - 2\cdot\nu_{21}\cdot\nu_{32}\cdot\nu_{13}}{E_1\cdot E_2\cdot E_3}}$$

$$7.7 = \frac{\nu_{31} + \nu_{21}\cdot\nu_{32}}{E_2\cdot E_3\cdot\dfrac{1 - \nu_{12}\cdot\nu_{21} - \nu_{23}\cdot\nu_{32} - \nu_{31}\cdot\nu_{13} - 2\cdot\nu_{21}\cdot\nu_{32}\cdot\nu_{13}}{E_1\cdot E_2\cdot E_3}}$$

$$31.5 = \frac{1 - \nu_{12}\cdot\nu_{21}}{E_1\cdot E_2\cdot\dfrac{1 - \nu_{12}\cdot\nu_{21} - \nu_{23}\cdot\nu_{32} - \nu_{31}\cdot\nu_{13} - 2\cdot\nu_{21}\cdot\nu_{32}\cdot\nu_{13}}{E_1\cdot E_2\cdot E_3}}$$

$$\begin{pmatrix} \nu_{21val} \\ \nu_{31val} \\ \nu_{23val} \\ \nu_{12val} \\ \nu_{13val} \\ \nu_{32val} \end{pmatrix} := \mathrm{Find}(\nu_{21}, \nu_{31}, \nu_{23}, \nu_{12}, \nu_{13}, \nu_{32}) \qquad \begin{pmatrix} \nu_{21val} \\ \nu_{31val} \\ \nu_{23val} \\ \nu_{12val} \\ \nu_{13val} \\ \nu_{32val} \end{pmatrix} = \begin{pmatrix} 0.092 \\ 0.091 \\ 0.222 \\ 0.092 \\ 0.222 \\ 0.091 \end{pmatrix}$$

Reciprocal relations check

$$\frac{\nu_{21val}}{E_2} = 1.264\times 10^{-3}$$

$$\frac{\nu_{12val}}{E_1} = 1.264\times 10^{-3}$$

$$\frac{\nu_{31val}}{E_3} = 3.033\times 10^{-3} \qquad \frac{\nu_{32val}}{E_3} = 3.033\times 10^{-3}$$

$$\frac{\nu_{13val}}{E_1} = 3.033\times 10^{-3} \qquad \frac{\nu_{23val}}{E_2} = 3.033\times 10^{-3}$$

Resulting Engineering Constants @ Low Temperature

$E_1 = 73.138$

$E_2 = 73.138$

$E_3 = 30.094$

$G_{23} := \frac{1}{S_{44}}$ $G_{23} = 11.3$

$G_{31} := \frac{1}{S_{55}}$ $G_{31} = 11.3$

$G_{12} := \frac{1}{S_{66}}$ $G_{12} = 19.6$

$\nu_{21} := \nu_{21val}$ $\nu_{21} = 0.092$

$\nu_{31} := \nu_{31val}$ $\nu_{31} = 0.091$

$\nu_{23} := \nu_{23val}$ $\nu_{23} = 0.222$

$\sigma_1 := .0758 \quad \sigma_2 := \sigma_1 \quad \sigma_3 := \left(\frac{2}{3}\sigma_1\right)$

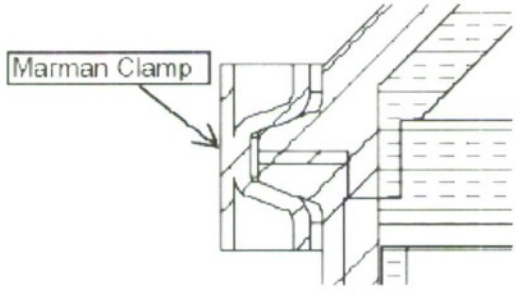
$\tau_{12} := .062 \quad \tau_{31} := \tau_{12} \quad \tau_{23} := \tau_{12}$

$$\begin{pmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \\ \tau_{23} \\ \tau_{31} \\ \tau_{12} \end{pmatrix} = \begin{pmatrix} 0.076 \\ 0.076 \\ 0.051 \\ 0.062 \\ 0.062 \\ 0.062 \end{pmatrix} \qquad \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \gamma_{23} \\ \gamma_{31} \\ \gamma_{12} \end{pmatrix} := \begin{pmatrix} S_{11} & S_{12} & S_{13} & S_{14} & S_{15} & S_{16} \\ S_{21} & S_{22} & S_{23} & S_{24} & S_{25} & S_{26} \\ S_{31} & S_{32} & S_{33} & S_{34} & S_{35} & S_{36} \\ S_{41} & S_{42} & S_{43} & S_{44} & S_{45} & S_{46} \\ S_{51} & S_{52} & S_{53} & S_{54} & S_{55} & S_{56} \\ S_{61} & S_{62} & S_{63} & S_{64} & S_{65} & S_{66} \end{pmatrix} \cdot \begin{pmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \\ \tau_{23} \\ \tau_{31} \\ \tau_{12} \end{pmatrix}$$

$$\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \gamma_{23} \\ \gamma_{31} \\ \gamma_{12} \end{pmatrix} = \begin{pmatrix} 7.873 \times 10^{-4} \\ 7.873 \times 10^{-4} \\ 1.219 \times 10^{-3} \\ 5.487 \times 10^{-3} \\ 5.487 \times 10^{-3} \\ 3.163 \times 10^{-3} \end{pmatrix}$$

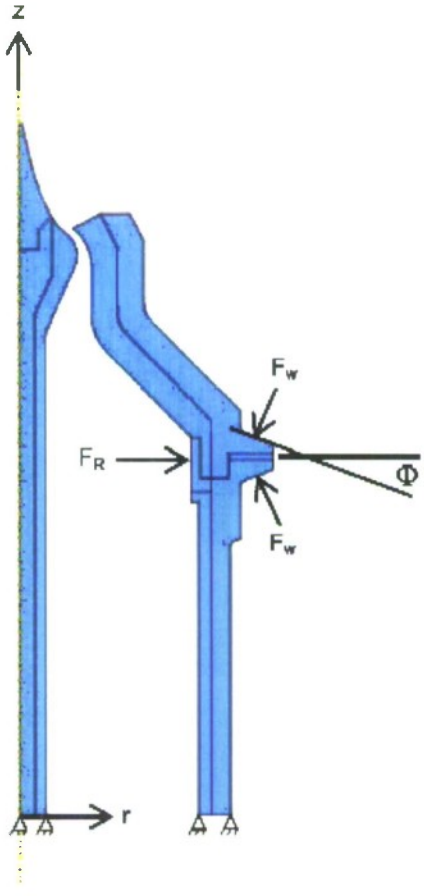
APPENDIX H

Marman Clamp Load: Sample Calculations



$$F_w = \frac{F_R}{2 \cdot \sin\phi}$$

$\phi := 20\text{deg}$



Reduce the assembly to a pressure vessel with a local band clamp

If pressure vessel is used

Then P=internal pressure

And $P = 2 \cdot \frac{N_\theta}{R_\theta}$ (from Adv. Mechanics of Materials[35] for a cylinder; membrane forces and stresses.)

where

N_θ is the tangential line load

R_θ is the radius of the cylinder

For a hose clamp(band clamp)

$$P = \frac{F_R}{2 \cdot \pi R_\theta \cdot w}$$

where

F_R is the radial clamping force

w width of the clamp(band)

Also,

$$N_\theta = \frac{F_T}{w}$$

where

F_T is the tangential clamping force

And so, by combining expression for the pressure vessel and the band clamp:

$$\frac{F_R}{2 \cdot \pi \cdot R_\theta \cdot w} = \frac{2 \cdot F_T \cdot \pi}{w \cdot R_\theta}$$

$$F_R = 4 \cdot F_T \cdot \pi$$

Need and expression for the tangential force in terms of the torque applied to the Marman clamp tightening screw[32].

$$F_T = \frac{T}{k \cdot d}$$

where

$T := 100 lbf \cdot in$ manufacturers torque spec. on the bolt/screw

$k := 0.2$ for bolt diameters ranging from .25" to 1"

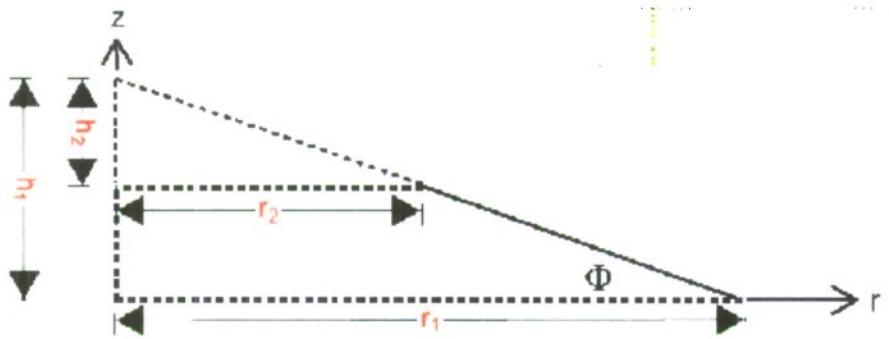
$d := 0.25 in$ and is the bolt diameter for the clamp

$$F_T := \frac{T}{k \cdot d} \qquad F_T = 8.896 \times 10^3 \, N$$

$$F_R := 4 \cdot F_T \cdot \pi \qquad F_R = 1.118 \times 10^5 \, N$$

$$F_w := \frac{F_R}{2 \cdot \sin(\phi)} \qquad F_w = 1.634 \times 10^5 \, N$$

Now the force needs to be divided by the approximate surface to which the Marman clamp is applied. This requires calculating a the surface area for the section of a cone



For the Marman clamp used for the assembly(Aeroquip AEB 258, part number 4575-400):

$r_1 := .123m$ $r_2 := .116m$

$h_1 := r_1 \cdot \tan(\phi)$ $h_2 := r_2 \cdot \tan(\phi)$

$$A_m := \pi \cdot r_1 \cdot \left(\sqrt{r_1^2 + h_1^2}\right) - \pi \cdot r_2 \cdot \left(\sqrt{r_2^2 + h_2^2}\right)$$

$$A_m = 5.593 \times 10^{-3} m^2$$

The pressure acting on the assembly by the Marman clamp is then:

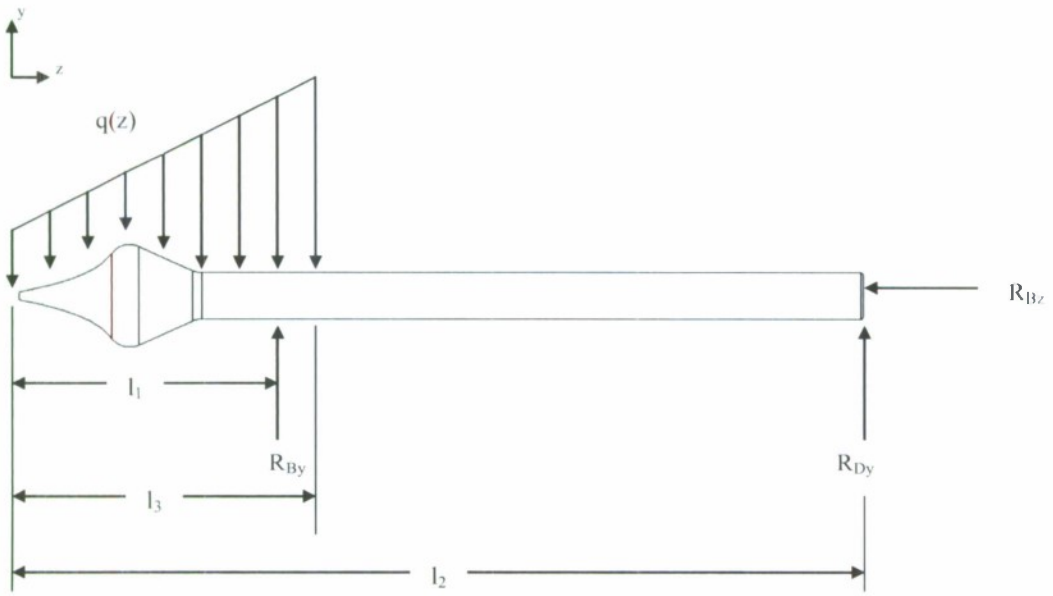
$$P_w := \frac{F_w}{A_m}$$

$P_w = 29.22MPa$

APPENDIX I

Spike Web Lateral Loading: Sample Calculations

The foregoing is a detailed calculation of the lateral load applied to the central hole of the spike web. Although the loading case seems unrealistic, it provides a conservative approach to assessing the failure of this part.



Operating pressures: TEP was used to determine the pressure at the tip of the spike if the chamber pressure is set at 450psi. The pressure distribution is assume to be linearly along the length of the spike, but is is not neccessarily physically accurate.

Chamber	Throat
$P_c := 450psi$	$P_t := 22.5psi$

Axial locations:

$l_1 := .07899927m$ Distance from the tip to the center plane of the spike web; location of R_By.

$l_2 := 0.23301696m$ Distance from the tip to the bottom of the combustion chamber(the entire length of the beam); location of R_Dy

$l_3 := l_1 + .00635m$ Distance from the throat to the end of the spike web; the load will be distributed along this length.

l_4 = unknown Distance from the tip to the point where an equivalent load is applied to represent the distributed load in the moment equation for static equilibrium; it will be solved for when satisfying the static equilibrium.

Sliced area of the spike:

$A_s := .00122323m^2$ Area of the planar surface that results when the spike is cut along the x-z plane.

Representing the distributed pressure as a distributed load:

$M := \frac{P_c - P_t}{l_3}$ The slope of the linearly distributed pressure.

$B := P_t$ The intercept

$P(z) := M \cdot z + B$ The linear equation for the distributed pressure

$q(z) := \frac{P(z) \cdot A_s}{l_3}$ The pressure is multiplied by the area to yield a load. The load is then distributed along the length l_3.

$q_{min} := q(0)$ $q_{min} = 2.223 \times 10^3 \frac{N}{m}$ Distributed load value at z=0

$q_{max} := q(l_3)$ $q_{max} = 4.447 \times 10^4 \frac{N}{m}$ Distributed load value at z=l_3

Determining the concentrated load:

$F_c := \int_0^{l_3} q(z)\, dz$ $F_c = 1.993 \times 10^3 N$ The change in shear along the beam from 0 to l_3 is equal to the integral of the distributed load equation.

Static Equilibrium

Sum of the forces in the z-direction: $R_{Dz} + R_{Bz} = 0$

Sum of the forces in the y-direction: $R_{By} + R_{Dy} - F_c = 0$

Sum of the moment about A:(the point at which the equilvalent concentrated force is applied will be solved for here).

$-R_{By} \cdot l_1 - R_{Dy} \cdot l_2 + F_c \cdot l_4 = 0$

The linearly distributed load(trapezoid-shaped) is broken up into a uniformly distributed load and a linearly distributed load(triangular-shaped) such that:

$$-R_{By} \cdot l_1 - R_{Dy} \cdot l_2 + F_c \cdot l_4 = -R_{By} \cdot l_1 - R_{Dy} \cdot l_2 + q_{min} \cdot l_3 \cdot \frac{l_3}{2} + (q_{max} - q_{min}) \cdot \frac{l_3}{2} \cdot \frac{2 \cdot l_3}{3}$$

Reducing the expression and dividing by F_c gives:

$$l_4 := \frac{q_{min} \cdot l_3 \cdot \frac{l_3}{2} + (q_{max} - q_{min}) \cdot \frac{l_3}{2} \cdot \frac{2 \cdot l_3}{3}}{F_c} \qquad l_4 = 0.056m$$

And so:

$-R_{By} \cdot l_1 - (R_{By} + F_c) \cdot l_2 + F_c \cdot l_4 = 0$

$-R_{By} \cdot l_1 + R_{By} \cdot l_2 - F_c \cdot l_2 + F_c \cdot l_4 = 0$

$R_{By} := \frac{F_c \cdot (l_2 - l_4)}{l_2 - l_1}$ $R_{By} = 2.296 \times 10^3 N$

$R_{Dy} := F_c - R_{By}$ $R_{Dy} = -303.427N$

The reaction force at B will be applied to an FEA model to assess failure

APPENDIX J

Spike Web FEA Results: Sample Calculations

FEA Beam model analysis

The results from above are used in two FEA models to determine the potential stresses seen by the spokes of the spike web. The model differ by the direction of the load applied. The figure below shpows the modified spike web model that includes internal beam elements to mimic the cross section of the aerospike shank.

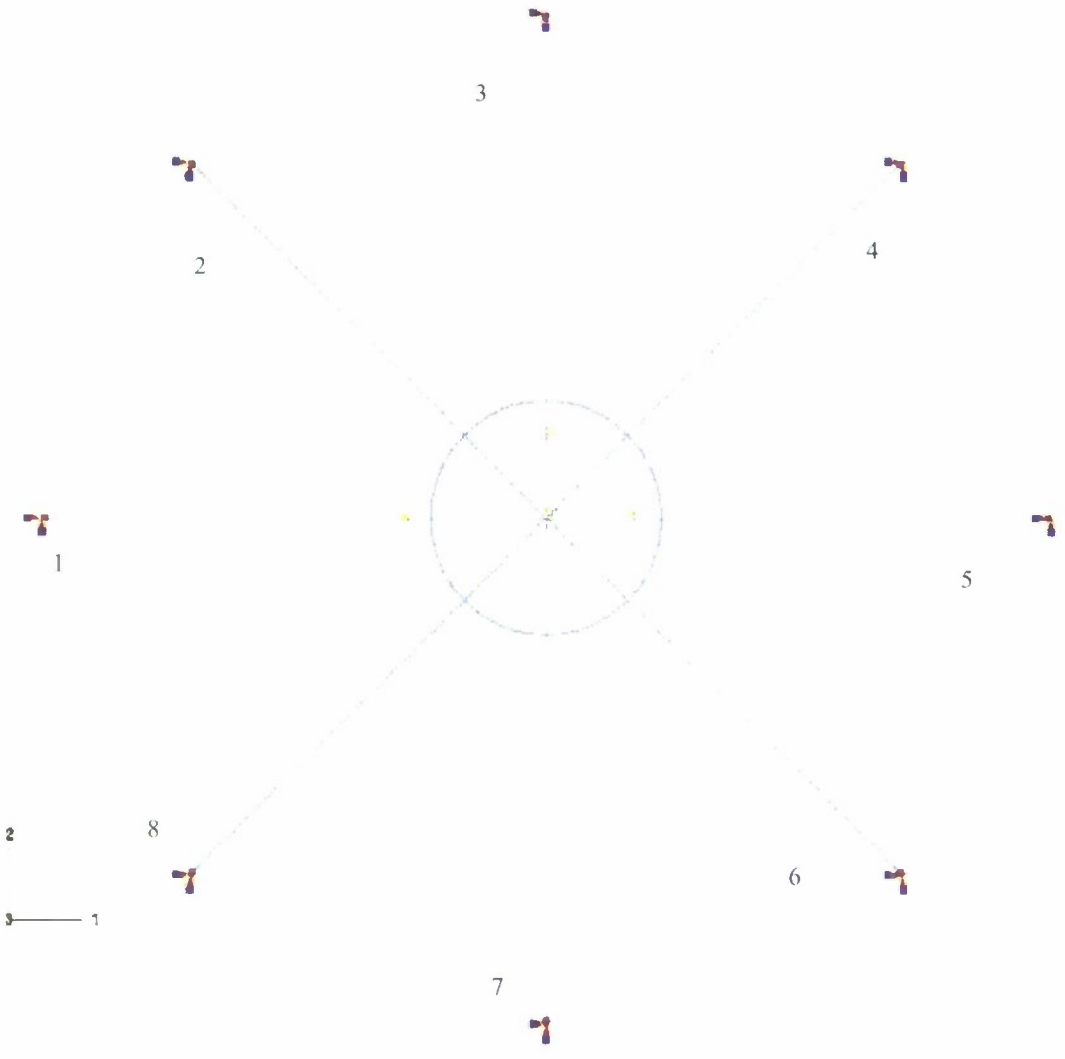


Figure LC1. The above model represents a modified version of previous models.

Load Case 1

Load Case 1 is repeated in the same manner.

Section dimensions

$a := .0127m$ This is the longest dimension in the FEA model section properties

$b := .003175m$ This is the shortest dimension in the FEA model section properties

$A_{cs} := a \cdot b$ $A_{cs} = 4.032 \times 10^{-5} m^2$ Cross sectional area

$I := \frac{1}{12} \cdot a \cdot b^3$ smallest moment of inertia

$I = 3.387 \times 10^{-11} m^4$

Length

$l := .0282579m$ span of the beam

Material Properties

$E := 11.8GPa$

Spoke #1

Section force

Compression	Tension	
$F_{sc} := 753.918N$	$F_{st} := 0N$	Pulled directly from the FEA results

This is actual unreasonable because of the interior beams that were added. The interior beam that is colinear to spoke #1 actual pulls back this spoke there by lessening the compression. In reality the load value is probably closer to the previous model geomtry.

The maximum bending moment for the spoke

$M_{max} := 0N \cdot m$ Pulled directly from the FEA results

Bending stress with regard to the above bending moment

$\sigma_b := \frac{M_{max} \cdot \frac{b}{2}}{I}$ $\sigma_b = 0MPa$ $\sigma_b = 0ksi$

The resulting compressive and tensile stresses for the two beams that are colinear with the applied load are as follows:

$\sigma_c := \frac{F_{sc}}{A_{cs}}$ $\sigma_c = 18.697MPa$ $\sigma_c = 2.712ksi$

$\sigma_t := \frac{F_{st}}{A_{cs}}$ $\sigma_t = 0MPa$ $\sigma_t = 0ksi$

Total compressive stress:

$\sigma_{ctot} := \sigma_c + \sigma_b$ $\sigma_{ctot} = 18.697\text{MPa}$

Total tensile stress:

$\sigma_{ttot} := \sigma_b + \sigma_t$ $\sigma_{ttot} = 0\,\text{MPa}$

Ultimate Strengths Values For SiC-6 Graphite:

Compressive	Tensile	Flexural
$\sigma_{uc} := 103\text{MPa}$	$\sigma_{ut} := 29\text{MPa}$	$\sigma_{uf} := 49\text{MPa}$

Compressive failure index:

$FI_c := \frac{|\sigma_{ctot}|}{\sigma_{uc}}$ $FI_c = 0.182$

Tensile failure index:

$FI_t := \frac{|\sigma_{ttot}|}{\sigma_{ut}}$ $FI_t = 0$

Bending failure index:

$FI_b := \frac{|\sigma_b|}{\sigma_{ut}}$ $FI_b = 0$

Spoke #2

Section force

Compression	Tension	
$F_{max} := 367.031\text{N}$	$F_{min} := 0\text{N}$	Pulled directly from the FEA results

The maximum bending moment for the spoke

$M_{max} := .0435399\text{N}\cdot\text{m}$ Pulled directly from the FEA results

Bending stress with regard to the above bending moment

$\sigma_b := \frac{M_{max}\cdot\frac{b}{2}}{I}$ $\sigma_b = 2.041\text{MPa}$ $\sigma_b = 0.296\text{ksi}$

The resulting compressive and tensile stresses for the two beams that are colinear with the applied load are as follows:

$\sigma_c := \frac{F_{sc}}{A_{cs}}$ $\sigma_c = 9.102\text{MPa}$ $\sigma_c = 1.32\text{ksi}$

$\sigma_t := \frac{F_{st}}{A_{cs}}$ $\sigma_t = 0\text{MPa}$ $\sigma_t = 0\text{ksi}$

Total compressive stress:

$\sigma_{ctot} := \sigma_c + \sigma_b$ $\sigma_{ctot} = 11.143\text{MPa}$

Total tensile stress:

$\sigma_{ttot} := \sigma_b + \sigma_t$ $\sigma_{ttot} = 2.041\text{MPa}$

Ultimate Strengths Values For SiC-6 Graphite:

Compressive	Tensile	Flexural
$\sigma_{uc} := 103\text{MPa}$	$\sigma_{ut} := 29\text{MPa}$	$\sigma_{uf} := 49\text{MPa}$

Compressive failure index:

$FI_c := \frac{|\sigma_{ctot}|}{\sigma_{uc}}$ $FI_c = 0.108$

Tensile failure index:

$FI_t := \frac{|\sigma_{ttot}|}{\sigma_{ut}}$ $FI_t = 0.07$

Bending failure index:

$FI_b := \frac{|\sigma_b|}{\sigma_{ut}}$ $FI_b = 0.07$

Spoke #3

Section force

Compression Tension

$F_{sc} := 0N$ $F_{st} := 20.5886N$ Pulled directly from the FEA results

The maximum bending moment for the spoke

$M_{max} := .0973399N \cdot m$ Pulled directly from the FEA results

Bending stress with regard to the above bending moment

$$\sigma_b := \frac{M_{max}\frac{b}{2}}{I}$$ $\sigma_b = 4.562MPa$ $\sigma_b = 0.662ksi$

The resulting compressive and tensile stresses for the two beams that are colinear with the applied load are as follows:

$$\sigma_c := \frac{F_{sc}}{A_{cs}}$$ $\sigma_c = 0MPa$ $\sigma_c = 0ksi$

$$\sigma_t := \frac{F_{st}}{A_{cs}}$$ $\sigma_t = 0.511MPa$ $\sigma_t = 0.074ksi$

Total compressive stress:

$\sigma_{ctot} := \sigma_c + \sigma_b$ $\sigma_{ctot} = 4.562MPa$

Total tensile stress:

$\sigma_{ttot} := \sigma_b + \sigma_t$ $\sigma_{ttot} = 5.073MPa$

Ultimate Strengths Values For SiC-6 Graphite:

Compressive Tensile Flexural

$\sigma_{uc} := 103MPa$ $\sigma_{ut} := 29MPa$ $\sigma_{uf} := 49MPa$

Compressive failure index:

$$FI_c := \frac{|\sigma_{ctot}|}{\sigma_{uc}}$$ $FI_c = 0.044$

Tensile failure index:

$$FI_t := \frac{|\sigma_{ttot}|}{\sigma_{ut}}$$ $FI_t = 0.175$

Bending failure index:

$$FI_b := \frac{|\sigma_b|}{\sigma_{ut}}$$ $FI_b = 0.157$

Spoke #4

Section force

Compression Tension

$F_{sc} := 0N$ $F_{st} := 355.795N$ Pulled directly from the FEA results

The maximum bending moment for the spoke

$M_{max} := .0678158N \cdot m$ Pulled directly from the FEA results

Bending stress with regard to the above bending moment

$$\sigma_b := \frac{M_{max} \cdot \frac{b}{2}}{I}$$ $\sigma_b = 3.178MPa$ $\sigma_b = 0.461ksi$

The resulting compressive and tensile stresses for the two beams that are colinear with the applied load are as follows:

$$\sigma_c := \frac{F_{sc}}{A_{cs}}$$ $\sigma_c = 0MPa$ $\sigma_c = 0ksi$

$$\sigma_t := \frac{F_{st}}{A_{cs}}$$ $\sigma_t = 8.824MPa$ $\sigma_t = 1.28ksi$

Total compressive stress:

$\sigma_{ctot} := \sigma_c + \sigma_b$ $\sigma_{ctot} = 3.178MPa$

Total tensile stress:

$\sigma_{ttot} := \sigma_b + \sigma_t$ $\sigma_{ttot} = 12.002MPa$

Ultimate Strengths Values For SiC-6 Graphite:

Compressive Tensile Flexural

$\sigma_{uc} := 103MPa$ $\sigma_{ut} := 29MPa$ $\sigma_{uf} := 49MPa$

Compressive failure index:

$$FI_c := \frac{|\sigma_{ctot}|}{\sigma_{uc}}$$ $FI_c = 0.031$

Tensile failure index:

$$FI_t := \frac{|\sigma_{ttot}|}{\sigma_{ut}}$$ $FI_t = 0.414$

Bending failure index:

$$FI_b := \frac{|\sigma_b|}{\sigma_{ut}}$$ $FI_b = 0.11$

Spoke #5

Section force

Compression Tension

$F_{sc} := 0N$ $F_{st} := 494.738N$ Pulled directly from the FEA results

The maximum bending moment for the spoke

$M_{max} := 0N \cdot m$ Pulled directly from the FEA results

Bending stress with regard to the above bending moment

$$\sigma_b := \frac{M_{max} \cdot \frac{b}{2}}{I}$$ $\sigma_b = 0MPa$ $\sigma_b = 0ksi$

The resulting compressive and tensile stresses for the two beams that are colinear with the applied load are as follows:

$$\sigma_c := \frac{F_{sc}}{A_{cs}}$$ $\sigma_c = 0MPa$ $\sigma_c = 0ksi$

$$\sigma_t := \frac{F_{st}}{A_{cs}}$$ $\sigma_t = 12.27MPa$ $\sigma_t = 1.78ksi$

Total compressive stress:

$\sigma_{ctot} := \sigma_c + \sigma_b$ $\sigma_{ctot} = 0MPa$

Total tensile stress:

$\sigma_{ttot} := \sigma_b + \sigma_t$ $\sigma_{ttot} = 12.27MPa$

Ultimate Strengths Values For SiC-6 Graphite:

Compressive Tensile Flexural

$\sigma_{uc} := 103MPa$ $\sigma_{ut} := 29MPa$ $\sigma_{uf} := 49MPa$

Compressive failure index:

$$FI_c := \frac{|\sigma_{ctot}|}{\sigma_{uc}}$$ $FI_c = 0$

Tensile failure index:

$$FI_t := \frac{|\sigma_{ttot}|}{\sigma_{ut}}$$ $FI_t = 0.423$

Bending failure index:

$$FI_b := \frac{|\sigma_b|}{\sigma_{ut}}$$ $FI_b = 0$

Load Case 2

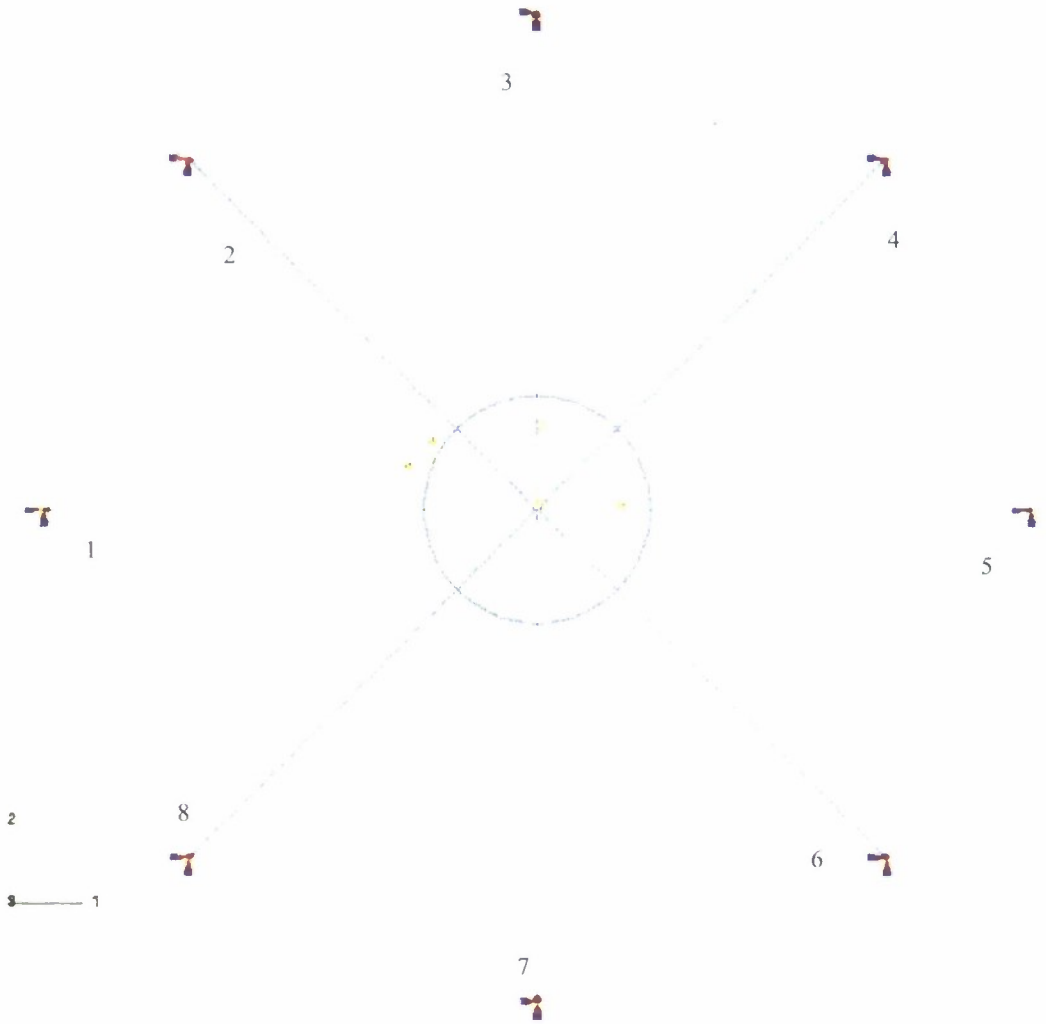


Figure LC2. The above model represents a modified version of previous models.

Section dimensions

$a := .0127m$ This is the longest dimension in the FEA model section properties

$b := .003175m$ This is the shortest dimension in the FEA model section properties

$A_{cs} := a \cdot b$ $A_{cs} = 4.032 \times 10^{-5} m^2$ Cross sectional area

$I := \frac{1}{12} \cdot a \cdot b^3$ smallest moment of inertia

Length

$l := .0282579m$ span of the beam

Material Properties

$E := 11.8GPa$

Spoke #1

Section force

Compression Tension

$F_{sc} := 606.639N$ $F_{st} := 0N$ Pulled directly from the FEA results

The maximum bending moment for the spoke

$M_{max} := .0352259N \cdot m$ Pulled directly from the FEA results

Bending stress with regard to the above bending moment

$$\sigma_b := \frac{M_{max} \cdot \frac{b}{2}}{I}$$ $\sigma_b = 1.651MPa$ $\sigma_b = 0.239ksi$

The resulting compressive and tensile stresses for the two beams that are colinear with the applied load are as follows:

$$\sigma_c := \frac{F_{sc}}{A_{cs}}$$ $\sigma_c = 15.045MPa$ $\sigma_c = 2.182ksi$

$$\sigma_t := \frac{F_{st}}{A_{cs}}$$ $\sigma_t = 0MPa$ $\sigma_t = 0ksi$

Total compressive stress:

$\sigma_{ctot} := \sigma_c + \sigma_b$ $\sigma_{ctot} = 16.696MPa$

Total tensile stress:

$\sigma_{ttot} := \sigma_b + \sigma_t$ $\sigma_{ttot} = 1.651MPa$

Ultimate Strengths Values For SiC-6 Graphite:

Compressive Tensile Flexural

$\sigma_{uc} := 103MPa$ $\sigma_{ut} := 29MPa$ $\sigma_{uf} := 49MPa$

Compressive failure index:

$$FI_c := \frac{|\sigma_{ctot}|}{\sigma_{uc}}$$ $FI_c = 0.162$

Tensile failure index:

$$FI_t := \frac{|\sigma_{ttot}|}{\sigma_{ut}}$$ $FI_t = 0.057$

Bending failure index:

$$FI_b := \frac{|\sigma_b|}{\sigma_{ut}}$$ $FI_b = 0.057$

Spoke #2 Because of symmetry spoke#2 has the same results as spoke#1

Spoke #3

Section force

Compression Tension

$F_{sc} := 168.149N$ $F_{st} := 0N$ Pulled directly from the FEA results

The maximum bending moment for the spoke

$M_{max} := .0981713N \cdot m$ Pulled directly from the FEA results

Bending stress with regard to the above bending moment

$$\sigma_b := \frac{M_{max} \frac{b}{2}}{1}$$

$\sigma_b = 4.601MPa$ $\sigma_b = 0.667ksi$

The resulting compressive and tensile stresses for the two beams that are colinear with the applied load are as follows:

$$\sigma_c := \frac{F_{sc}}{A_{cs}}$$

$\sigma_c = 4.17MPa$

$\sigma_c = 0.605ksi$

$$\sigma_t := \frac{F_{st}}{A_{cs}}$$

$\sigma_t = 0MPa$

$\sigma_t = 0ksi$

Total compressive stress:

$\sigma_{ctot} := \sigma_c + \sigma_b$ $\sigma_{ctot} = 8.771MPa$

Total tensile stress:

$\sigma_{ttot} := \sigma_b + \sigma_t$ $\sigma_{ttot} = 4.601MPa$

Ultimate Strengths Values For SiC-6 Graphite:

Compressive Tensile Flexural

$\sigma_{uc} := 103MPa$ $\sigma_{ut} := 29MPa$ $\sigma_{uf} := 49MPa$

Compressive failure index:

$$FI_c := \frac{|\sigma_{ctot}|}{\sigma_{uc}}$$

$FI_c = 0.085$

Tensile failure index:

$$FI_t := \frac{|\sigma_{ttot}|}{\sigma_{ut}}$$

$FI_t = 0.159$

Bending failure index:

$$FI_b := \frac{|\sigma_b|}{\sigma_{ut}}$$

$FI_b = 0.159$

Spoke #4

Section force

Compression Tension

$F_{sc} := 0N$ $F_{st} := 209.009N$ Pulled directly from the FEA results

The maximum bending moment for the spoke

$M_{max} := .0999374N \cdot m$ Pulled directly from the FEA results

Bending stress with regard to the above bending moment

$$\sigma_b := \frac{M_{max} \cdot \frac{b}{2}}{I}$$

$\sigma_b = 4.684MPa$ $\sigma_b = 0.679ksi$

The resulting compressive and tensile stresses for the two beams that are colinear with the applied load are as follows:

$$\sigma_c := \frac{F_{sc}}{A_{cs}}$$

$\sigma_c = 0MPa$

$\sigma_c = 0ksi$

$$\sigma_t := \frac{F_{st}}{A_{cs}}$$

$\sigma_t = 5.183MPa$

$\sigma_t = 0.752ksi$

Total compressive stress:

$\sigma_{ctot} := \sigma_c + \sigma_b$ $\sigma_{ctot} = 4.684MPa$

Total tensile stress:

$\sigma_{ttot} := \sigma_b + \sigma_t$ $\sigma_{ttot} = 9.867MPa$

Ultimate Strengths Values For SiC-6 Graphite:

Compressive Tensile Flexural

$\sigma_{uc} := 103MPa$ $\sigma_{ut} := 29MPa$ $\sigma_{uf} := 49MPa$

Compressive failure index:

$$FI_c := \frac{|\sigma_{ctot}|}{\sigma_{uc}}$$

$FI_c = 0.045$

Tensile failure index:

$$FI_t := \frac{|\sigma_{ttot}|}{\sigma_{ut}}$$

$FI_t = 0.34$

Bending failure index:

$$FI_b := \frac{|\sigma_b|}{\sigma_{ut}}$$

$FI_b = 0.162$

Spoke #5

Section force

Compression Tension

$F_{sc} := 0N$ $F_{st} := 465.808N$ Pulled directly from the FEA results

The maximum bending moment for the spoke

$M_{max} := .0382962N \cdot m$ Pulled directly from the FEA results

Bending stress with regard to the above bending moment

$$\sigma_b := \frac{M_{max} \frac{b}{2}}{I}$$

$\sigma_b = 1.795MPa$ $\sigma_b = 0.26ksi$

The resulting compressive and tensile stresses for the two beams that are colinear with the applied load are as follows:

$$\sigma_c := \frac{F_{sc}}{A_{cs}}$$

$\sigma_c = 0MPa$

$\sigma_c = 0ksi$

$$\sigma_t := \frac{F_{st}}{A_{cs}}$$

$\sigma_t = 11.552MPa$

$\sigma_t = 1.675ksi$

Total compressive stress:

$\sigma_{ctot} := \sigma_c + \sigma_b$ $\sigma_{ctot} = 1.795MPa$

Total tensile stress:

$\sigma_{ttot} := \sigma_b + \sigma_t$ $\sigma_{ttot} = 13.347MPa$

Ultimate Strengths Values For SiC-6 Graphite:

Compressive Tensile Flexural

$\sigma_{uc} := 103MPa$ $\sigma_{ut} := 29MPa$ $\sigma_{uf} := 49MPa$

Compressive failure index:

$$FI_c := \frac{|\sigma_{ctot}|}{\sigma_{uc}}$$

$FI_c = 0.017$

Tensile failure index:

$$FI_t := \frac{|\sigma_{ttot}|}{\sigma_{ut}}$$

$FI_t = 0.46$

Bending failure index:

$$FI_b := \frac{|\sigma_b|}{\sigma_{ut}}$$

$FI_b = 0.062$

It is the intent of this analysis to treat one single spoke of the spike web as a column with the potential to buckle. A pinned-pinned connection represents the most conservative column buckling analysis and will be used here. A derivation of the pinned-pinned column buckling formulation has been included.

$$P_{cr} := \frac{\pi^2 \cdot E \cdot I}{l^2} \qquad \text{Let:} \quad F_{sca} := 1338N$$

$$P_{cr} = 4.94 \times 10^3\ N$$

$$P_{actual} := |F_{sc}|$$

Buckling index

$$FI_{buckling} := \frac{P_{actual}}{P_{cr}} \qquad FI_{buckling} = 0.271$$

Assessment of blockage area:

The percent increase in the flow area will be approximated using the virtual measuring tools in Solidworks 2006.

Radius of opening if no spike web existed:

$r_o := 0.040462206m$

$r_i := .00635m$

$A_{max} := \pi \cdot \left(r_o^2 - r_i^2 \right)$ maximum area possible without a spike web

Surface area of hardware directly in the flow stream

Original configuration

$$A_{orig} := 1.423 \times 10^{-3} m^2 - \left[\left(\frac{.086868m}{2} \right)^2 \cdot \pi - A_{max} \right] \qquad A_{orig} = 5.131 \times 10^{-4} m^2$$

Percent Blockage

$$PB_{orig} := \frac{A_{orig}}{A_{max}} \qquad PB_{orig} = 0.102$$

Modified configuration

$$A_{mod} := 1.804 \times 10^{-3} m^2 - \left[\left(\frac{.086868m}{2} \right)^2 \cdot \pi - A_{max} \right] \qquad A_{mod} = 8.941 \times 10^{-4} m^2$$

Percent Blockage

$$PB_{mod} := \frac{A_{mod}}{A_{max}} \qquad PB_{mod} = 0.178$$

7.6% increase in blockage of area

APPENDIX K

FEA Results: Supporting Abaqus Files

Sonrce Files:

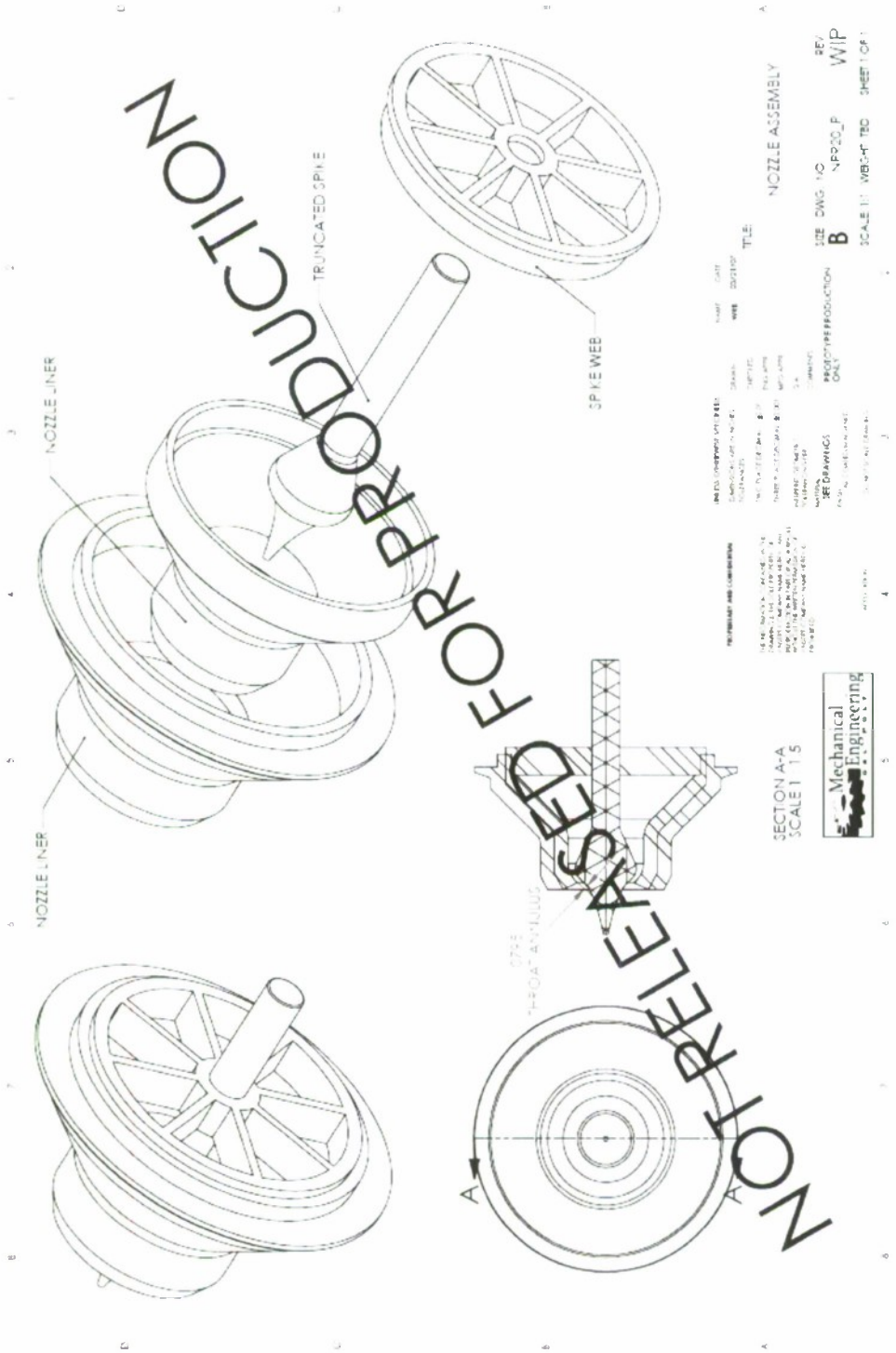
- ***Assembly Model*** *(various file formats with the following names)*
  - *NPR_20_HT_R5a through NPR_20_HT_R5j*
  - *NPR_20_ST_R5a, NPR_20_ST_R5d, NPR_20_ST_R5d_90 and NPR_20_ST_R5a*
- ***Spike Web Model*** *(various file formats with the following names)*
  - *NPR_20_HT_WEB*
  - *NPR_20_ST_WEB*

Notes:

- **.odb and *.inp file formats are available for the file names listed above.*
  - **.inp=input file*
  - **.odb=output database file*
- *In some case other supporting files exist*
- *HT=transient heat conduction model*
- *ST=thermal structural model*
- *R5=not important; model revision level*
- *a=spike configuration #1, b= spike configuration #2, etc.*
  - *refer to Table 16.2*
- *"_90"=thermal structural model with a 90° material rotation for the C/SiC*
- *WEB=spike web*

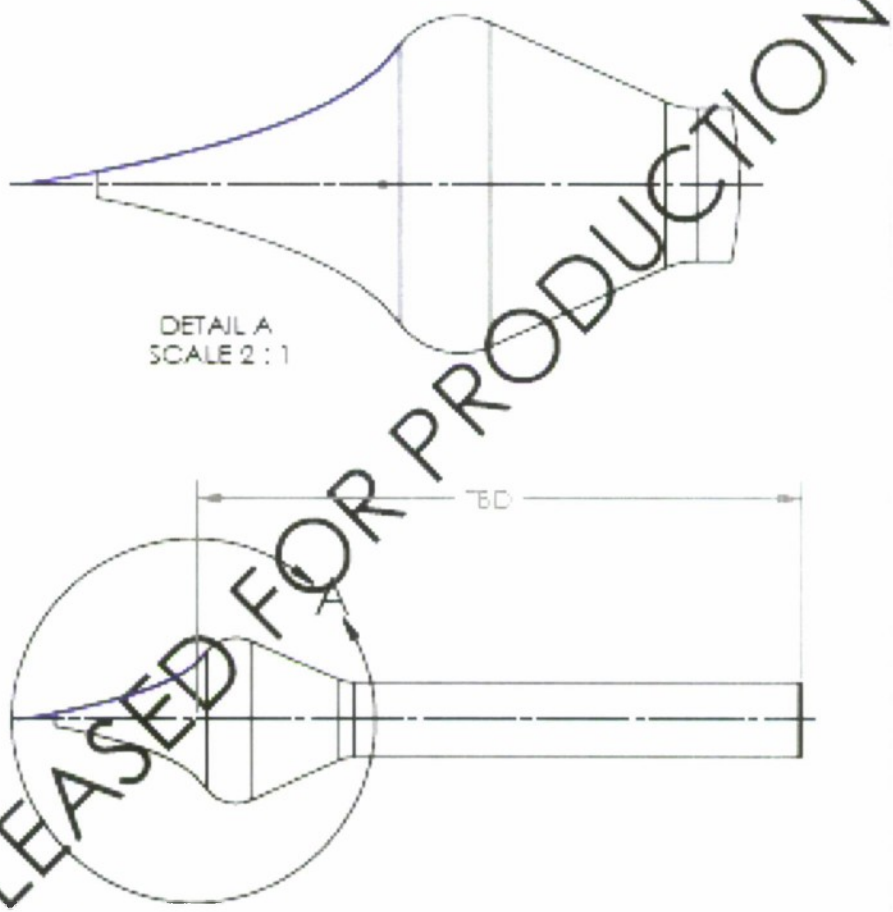
APPENDIX L

Working Drawings



NOT RELEASED FOR PRODUCTION
TRUNCATED SPIKE
SPIKE_NPR20_P
FOR REFERENCE
KNOCK DOWN EDGE
NOTES:
1. DIMENSIONS ARE GIVEN AS FINISHED.
Mechanical Engineering
PROTOTYPE PRODUCTION ONLY
TBD
R.25
48°

Y	Z
0.011635	-0.00149
0.011353	-0.00119
0.011093	-0.00106
0.010964	-0.00094
0.010828	-0.00084
0.01071	-0.00074
0.010598	-0.00064
0.010485	-0.00054
0.010375	-0.00044
0.010268	-0.00034
0.010157	-0.00024
0.010047	-0.00013
0.009936	-1.8E-05
0.009824	9.68E-05
0.00971	0.000217
0.009594	0.000342
0.009476	0.000473
0.009356	0.00061
0.009234	0.000756
0.009109	0.000909
0.008981	0.001068
0.00885	0.001234
0.008718	0.001411
0.008579	0.001595
0.008439	0.001795
0.008295	0.002006
0.008148	0.002226
0.007996	0.002459
0.007841	0.002707
0.007682	0.002969
0.007519	0.003246
0.007351	0.003541
0.007179	0.003853
0.007003	0.004185
0.006822	0.004537
0.006636	0.004911
0.006445	0.005309
0.006249	0.005732
0.006048	0.006183
0.005841	0.006661
0.005629	0.007
0.005411	0.007
0.005188	0.008
0.004958	0.008
0.004723	0.009
0.004481	0.010291
0.004233	0.01105
0.003978	0.011863
0.003717	0.012736
0.003449	0.01367
0.003174	0.014675
0.002891	0.015756
0.002602	0.01692
0.002304	0.018178
0.002	0.01963
0.001687	0.020956
0.001366	0.022584
0.001037	0.024306
0.0007	0.026179
0.000354	0.028219
0	0.030447



NOTES:

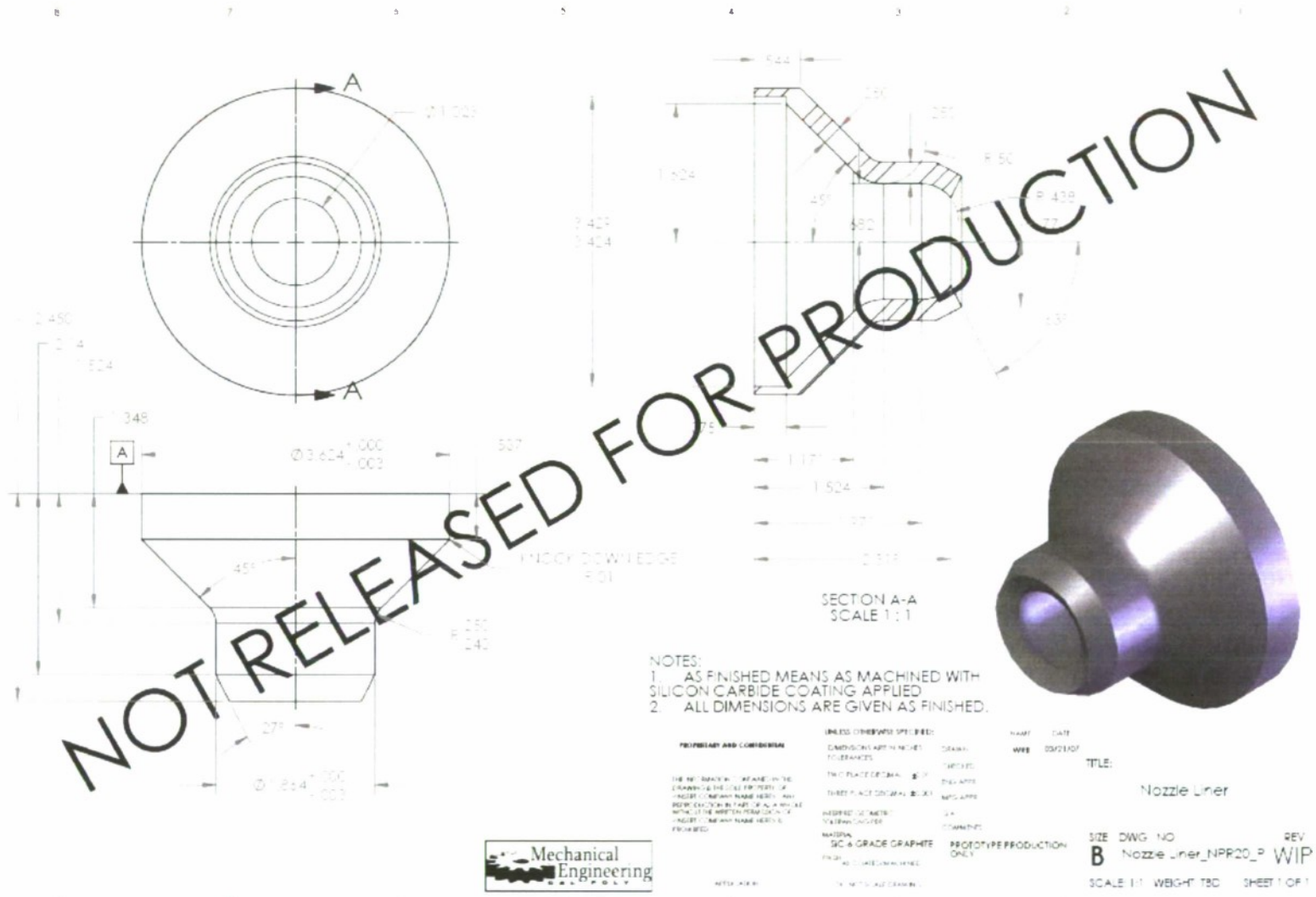
1. ALL Z AND Y COORDINATES GIVEN REPRESENT POINTS THAT MAKE UP THE SPLINE CURVE FOR THE SPIKE CURVATURE.
2. COORDINATE SYSTEM IS LOCATED AT THE CENTER OF THE OUTER ANNULUS RING OF THE ANNULUS AREA AT THE THROAT OF THE ASSEMBLED NOZZLE.
3. COORDINATES ARE GIVEN TO CREATE SPIKE CURVATURE BEYOND THE POINT OF TRUNCATATION. SPIKE SHOULD STILL BE TRUNCATED.
4. Z AND Y COORD. DATA IS GIVEN TO SIX DECIMAL PLACES, BUT A UNIFORM TOLERANCE OF ±0.01 ALONG THE LENGTH OF THE CURVE IS EXPECTED.

TRUNCATED SPIKE

A SPIKE NPR20 P

Mechanical Engineering

P.I.: Susan C. Opava, Ph.D.



NOT RELEASED FOR PRODUCTION
SECTION A-A
SCALE 1 : 1
DETAIL B
SCALE 3 : 1
Mechanical Engineering
Nozzle Outer
B Nozzle Outer_NP250_P_WIP

APPENDIX M

Material Properties

Source File:

Material Properties.xls

Notes:

1. Properties for various materials used in this study.

Attachment 2. Investigation of Heat Transfer in the Nozzle of Hybrid Rocket Engines

By

Eric Peterson

California State Polytechnic University

San Luis Obispo

2008

Nomenclature

Cp - Specifie Heat
D - Diameter
h - Heat Transfer Coeffieient
K - Thermal Conduetivity of Hot Gas
k - Thermal Conduetivity of Rod
Pr - Prandtl Number
Q - Heat Transfer
R - Reeovery Factor
Re - Reynold Number
St - Stanton Number
T_{ad} - Adiabatie Wall Temperure
T_g - Statie Temperure of Hot Gas
T_{wall} - Temperature of Wall Surface
V - Free Stream Gas Velocity

Executive Summary

Cal Poly is developing a hybrid-rocket engine with an aero-spike nozzle. In order to design a robust aero-spike nozzle it is important to know what the thermal boundary conditions will be at the throat of the nozzle. The throat is where the maximum amount of heat transfer will occur due to convection and radiation heat transfer. To accomplish this task my senior project will experimentally calculate the overall heat transfer coefficient due to radiation and convection of a standard converging-diverging nozzle as a starting point for further analysis.

Background

The hybrid rocket engine burns solid Plexiglas and liquid nitrous oxide at high temperatures to generate thrust. The rocket is ignited by heating the Plexiglas to 300 (F) by burning propane inside the chamber. Once the heat vaporizes the Plexiglas, nitrous oxide is injected into the combustion chamber and the combustion reaction begins.

The combustion process in similar to a burning candle[2]. The Plexiglas is vaporized by the heat of the flame. The gaseous Plexiglas flows into the flame zone inside the boundary layer. The Plexiglas fuel mixes with the nitrogen dioxide oxidizer inside the zone and undergoes a combustion reaction. The reaction produces hot exhaust gas that flows around 6000 F which then flow out through the rocket nozzle. Figure 1 shows the combustions process inside the rocket chamber upstream from the nozzle.

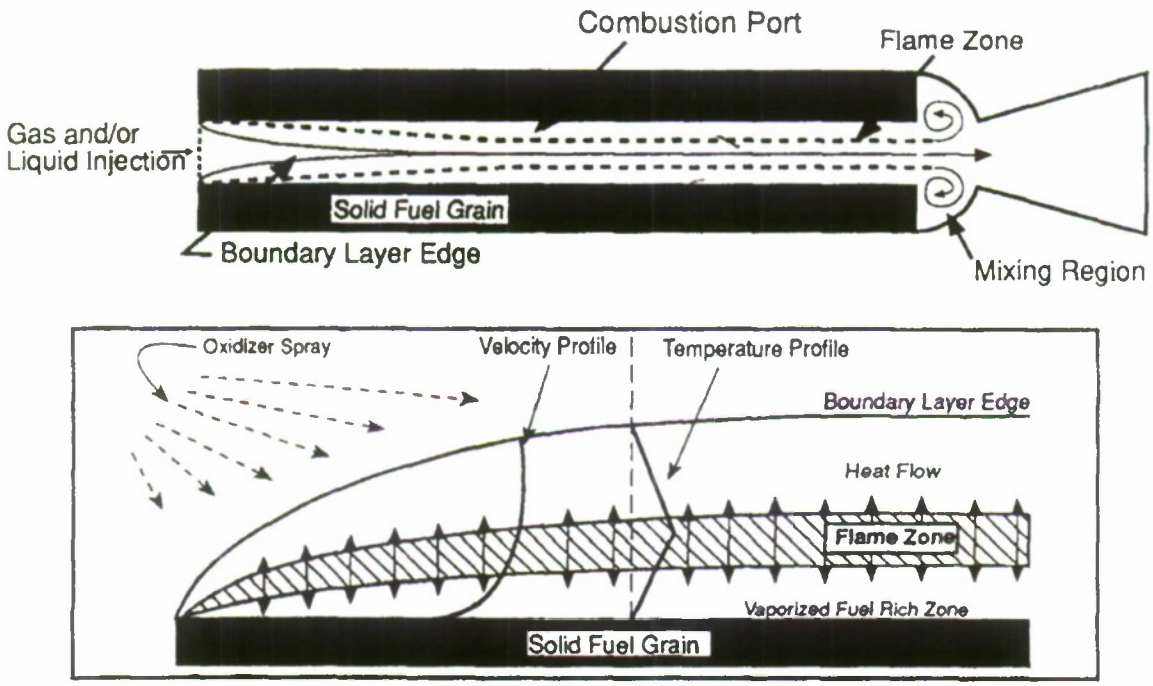


Figure (1) Combustion process for a hybrid rocket engine[1]

[2] An Evaluation of Scaling Effects For Hybrid Rocket Motors

The exhaust gases from the combustion process pass through the nozzle at high flow rates and undergo a rapid acceleration. In the converging section of the nozzle, leading up to the throat, the gases are accelerated to mock one. Once the gases reach mock one they past through the diverging section of the nozzle which continues to accelerate the flow. The properties of the gases change as they pass through the nozzle as a result of the rapid acceleration and compressibility of the gas. The area of focus for this project will be throat so it is important to know how the properties of the exhaust gas change leading up to the throat. In the converging section of the nozzle, the density, static temperature, static pressure, enthalpy, and cp of the exhaust gas decreases. The prandtl number increases. The gases inside the nozzle are turbulent due to its high velocity. Combustion can occur inside the nozzle resulting in flow distortions and changes in the gas temperature[3]. Oblique and normal shocks waves could occur in the throat of the nozzle. The properties of the exhaust gas will be calculated using a TEP program. A sherlin camera will be used to detect shock waves.

The velocity of the gas next to the wall must be zero due to the no slip condition resulting in a very large velocity gradient within the boundary layer. Heat is generated inside the boundary layer when faster moving molecules impact slower moving molecules resulting in a spike in the temperature of the gases. The maximum temperature inside the boundary layer is called the adiabatic wall temperature. The adiabatic wall temperature is used to calculate the heat flux at the surface of the nozzle instead of the free stream temperature[4]. This phenomenon can be seen in Figure 2. The thermal boundary layer is the regain of gas that exists between T_S and T_W seen in Figure 2.

[3] Advances in Heat Transfer
[4] Liquid-Propellant Rockets

Heat energy is transferred from the hot gas to the wall by convection and radiation heat transfer. The amount of heat transfer can be calculated using equation () where A is the cross sectional area of the rod, h is the overall heat transfer coefficient, T_{ad} is the adiabatic gas temperature, and T_{wall} is the temperature of the nozzle wall. The overall heat transfer coefficient (h) is a function of fluid velocity, fluid properties, nozzle shape, and throat diameter[5].

$$Q_{in}=A\ h\ (T_{ad}-T_{wall}) \qquad (1)$$

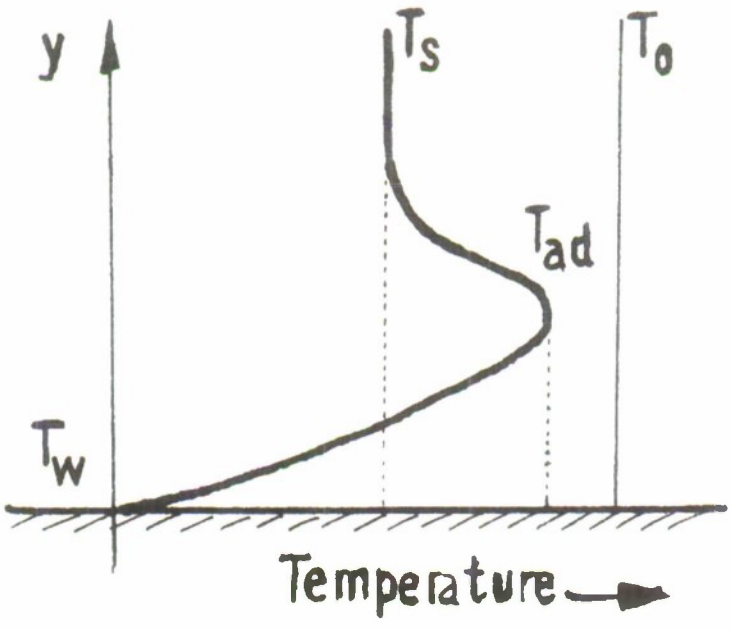


Figure (2) Temperature profile of fluid in the nozzle[6]

[5] Heat Transfer by Convection
[6] Heat Transfer in Rocket Engines

Heat Flux Sensor

Description

My senior project has two identical heat flux sensors inside a rocket nozzle. Each sensor measures the temperature profile of an inconel rod over time at three different points using 24gage K type thermal couples. A fourth thermal couple records the temperature of the air next to the inconel rod. The inconel rod is 0.126 inches in diameter. The throat of the nozzle is 0.5 inches in diameter. The inconel rods are place inside graphite nozzle with one end exposed to the hot exhaust gases in the throat and the other end against the wall of the stainless steel shell. The air gap around the inconel rod provides space for the thermal couples and the air acts as insulation. Both the air and graphite have low thermal conductivity relative to the inconel rods so 1-D heat transfer can be assumed. The graphite nozzle in housed inside a stainless steel shell which provides structural support. Conax wire glands screw into the stainless steel shell with Teflon tap around the screw threads. The thermal couple wires pass through the wire glands to the voltage recording instruments. Wire glands are used to create an air tight barrier between the air gap and the outside environment. The wire glands compress heat resistant powder around thermal couple wires to create an air tight barrier.

Diagram

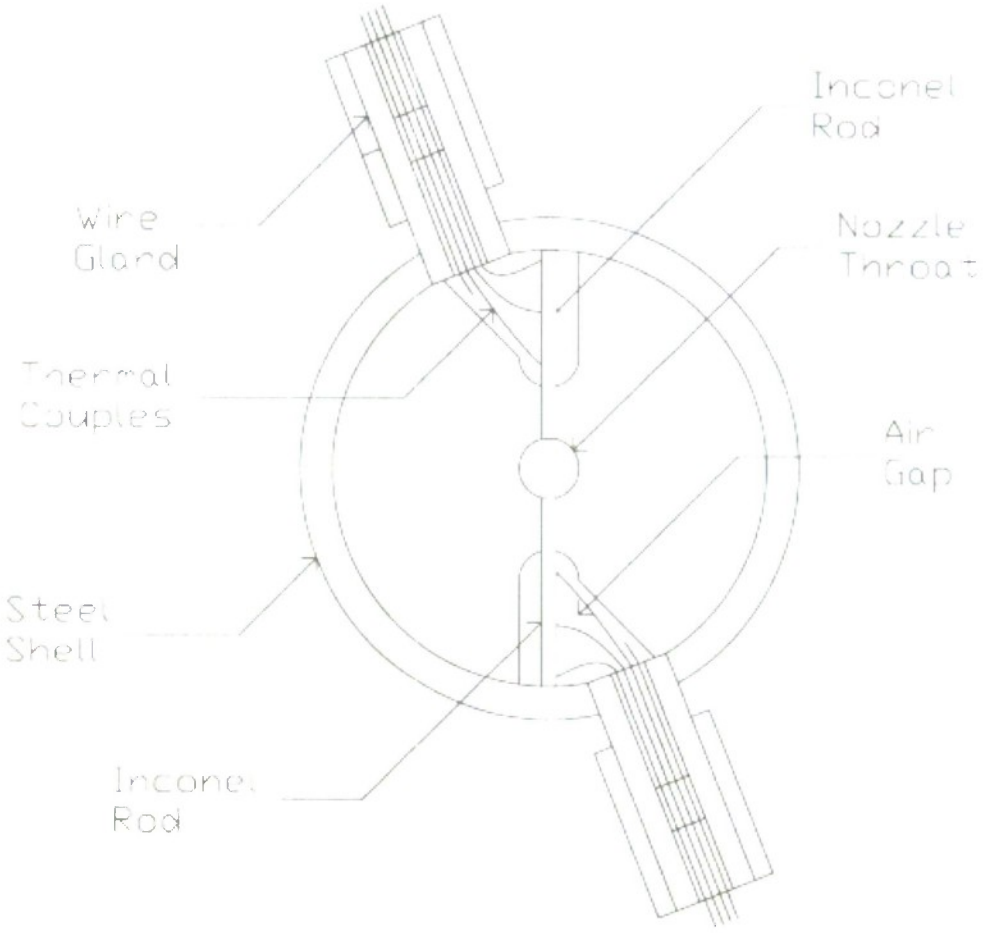


Figure (3) Diagram of heat flux sensor

Experiential Procedures

- Calibration
    - Before assembling the heat flux sensor the thermal couples need to be calibrated. The inconel rods need to be heated above room temperature to generate a voltage potential. The voltage potential at two different temperatures will be recorded and compared to the manufactures data to ensure accurate temperature recording.
- Assembly
    - First the nozzle is inserted in the steel shell.
    - Next the inconel rod is inserted into the nozzle through holes in the steel shell.
    - Then the thermal couple wires are pulled through the Conax wire gland.
    - Then the Conax wire glands are screwed into the steel shell.
    - Then the Conax wire gland is then sealed.
    - Finally the heat flux sensor is then attached the rocket chamber.
- Experiment
    - The first rocket burn will be at a low chamber pressure, 150 psi, for 4 seconds to test the seals of the heat flux sensor and thermal couple response.
    - If the thermal couple recording the air temperature inside the air gap records a spike in temperature during the 4 second burn, hot exhaust gas may have enter the air gap compromising the thermal couple data.
    - If temperature data looks accurate the rocket will be burned for 10 seconds at a chamber pressure 150 psi. The temperature of all four thermal couples will be recorded during the entire burn.
    - The thermal couple data from the second run will be transferred to excel for analysis.

Data Analysis

Once the temperature of the thermal couples over time is collected it will be easy to determine if steady state heat transfer was reached during the experiment. If the temperature readings of the thermal couples level off at a constant temperature and the slop of the temperature gradient along the rod is linear, a steady analysis of the data can be performed. If the temperature reading do not level off during the experiment a transient heat transfer analysis of the data must be performed.

Steady State Heat Transfer Analysis

The governing equation for steady state heat transfer shown in equation (2) which shows that the heat transfer entering the rod is equal to the heat transfer passing through the rod by conduction.

$$Q_{in} = Q_{conduction} \quad (2)$$

The Q_{in} is the heat transfer due to radiation and convection and is calculated using equation (3). The T_{wall}, the hot-side wall temperature, is approximated by assuming a linear temperature distribution in the rod. The thermal couple data is used as boundary condition to generate the linear distribution. The adiabatic wall temperature is calculated using equation (4) and (5) where T_g is the stagnation temperature of the hot gas.

$$Q_{in} = A\, h\, (T_{ad} - T_{wall}) \quad (3)$$

$$T_{ad} = T_g + RV^2/(2\,cp) \quad (4)$$

$$R = Pr^{1/3} \quad (5)$$

The $Q_{conduction}$ is the heat transfer passing through the rod where k is the thermal conductivity of the rod and A is the cross-sectional area. The value of dT/dx is a constant value along the entire length of the rod and is calculated using the thermal couple data.

$$Q_{conduction} = - k A (dT/dx) \quad (6)$$

Combining equation (3) and (6) results in equation (7). This is used to approximate the overall heat transfer coefficient.

$$h = -k(dT/dx)(1/(T_{ad}-T_{wall}) \quad (7)$$

Transient Heat Transfer Analysis

The transient analysis requires the use of curving fitting software to approximate the temperature profile of the entire rod at different point in time. Excel will provide T(x) for each time interval. The temperature profile will be non-linear. The T_{wall} will be approximated at each time interval using the temperature profile equations, $T(x_{wall})$. Two methods of calculating the overall heat transfer coefficient will be performed.

Method 1

The governing equation for transient heat transfer to the rod is shown in equation (8)

$$Q_{in} = Q_{conduction}(x_{wall}) \quad (8)$$

The amount of heat transfer entering the rod is calculated using equation (9) where $dT(x_{wall})/dx$ is the slop of the temperature profile at the end of the rod not at the thermal couples. Equation (10) is used to approximate overall heat transfer coefficient, h.

$$Q_{conduction} = - k A (dT(x_{wall})/dx) \quad (9)$$

$$h = -k(dT(x_{wall}/dx)(1/(T_{ad}-T_{wall}) \quad (10)$$

Method 2

The T_{wall} will continue to rise over time until it levels off at specific value. $T_{wall}(t)$ will can be approximated in excel. The steady state value of T_{wall} can be approximated by setting $t=\infty$. The steady state temperatures distribution in the rod can be calculated using the same technique. Once the steady state temperature profile, T(x), is known the constant value for dT/dx along the length the rod can be approximated. The heat transfer coefficient, h, can then be calculated using the steady state governing equation (7).

Comparing Results to Theoretical Values

The heat transfer coefficient can be calculated using three different methods. The first method uses the Stanton number correlations presented in equation (11). Stanton number can be assumed to be 0.002 from AGARD[4]. The h_g calculated using this method should be considered a rough approximation of the heat transfer coefficient.

$$h = St\ \rho\ V\ cp \tag{11}$$

The second method uses a Nussalt Number correlation presented in equation (12)[2]. When performing a 1-D analysis on a standard converging-diverging nozzle the internal diameter of the nozzle should be used. The K is the thermal conductivity of the fluid which is assumed to be approximately the thermal conductivity of N_2O. This method assumes the fluid properties are constant through out the boundary layer. The h_g calculated using this method should be considered a rough approximation of the heat transfer coefficient.

$$H = (K/D)\ Re^{0.8}\ Pr^{0.4} \tag{12}$$

The third method uses Bartz equation for rapidly calculating heat transfer in rocket nozzles presented in equation (13)[4]. When performing a 1-D analysis on a standard converging-diverging nozzle the internal diameter of the nozzle should be used. The variable σ is a property correlation factor which can be assumed to be 0.879. The remaining variables in the equation were taken from the TEP program provided by Dr Carpenter.

$$H = (0.026/D^{0.2})*(\mu^{0.2}*cp)/Pr^{0.6}\ (\rho*V)^{0.8}*\sigma_p \tag{13}$$

References

- D. R. Bartz. Advances in Heat Transfer. 1965.
- H. Ziebland & C. Parkinson. Heat Transfer in Rocket Engines, AGRAD. 1971.
- Liquid-Propellant Rockets
- P.N. Estey, D. Altman, & J.S. McFarlane. An Evaluation of Scaling Effects for Hybrid Rocket Motors. 1991.
- Ronald S. Mullisen. Heat Transfer by Convection. 1993.
- N. Sugathan, K. Srinivasan, & S. Srinivasa Murthy. Experiments on Heat Transfer in Cryogenic Engines Thrust Camber, Journal of Propulsion and Power, pg 240-244. 1993.
- A. Kumakawa, T. Tomita, H. Sakamoto. A Study of Plug-Nozzle Rocket Calorimeters. 1996.
- G. P. Sutton & O. Biblarz. Rocket Propulsion Elements. 2001.
- Kim Shollenberger
- Ronald S. Mullisen

Development of a Hybrid Rocket Motor Facility for Advanced Nozzle Research

Project Investigators:

William R. Murray and Thomas W. Carpenter
Mechanical Engineering
California Polytechnic State University
San Luis Obispo, CA

Development of a Hybrid Rocket Motor Facility for Advanced Nozzle Research,
William R. Murray and Thomas W. Carpenter

Project Objective

The objective of this project was to upgrade the existing Cal Poly hybrid rocket motor facility to state-of-the-art research standards. Now that this upgrade is completed, this facility is available for the testing of rocket nozzles (advanced converging-diverging nozzles, aerospike nozzles and advanced bell-shaped nozzles) under realistic hot-flow conditions. The propellant combination for our hybrid rocket motor is liquid nitrous oxide (N_2O) as the oxidizer and either Plexiglas® or hydroxyl-terminated polybutadiene (HTPB) as the fuel. In spite of the recent and still unexplained accident[1] at the Mojave Spaceport, N_2O-HTPB hybrid rocket motors are relatively safe because both the fuel and oxidizer can be stored without special precautions, and they do not burn when brought together without a significant source of heat for ignition. Not only are these compounds inherently safe and inexpensive, they are relatively non-polluting. To match the capabilities of this hybrid rocket motor facility to the capabilities of our new cold-flow blow-down facility being developed under separate funding[2], this hybrid rocket motor should be able to maintain a chamber pressure of 300 psia. Not only will this facility be beneficial for the testing of rocket nozzles, but the marriage of hybrid rocket motors and thrust vectored aerospike nozzles is interesting in its own right. The potential for hybrid rocket motors was clearly demonstrated in 2004, when Cal Poly alumni Burt Rutan's SpaceShipOne won the Ansari X-Prize using an N_2O-HTPB hybrid rocket to reach outer space.

Results

In the first year of this project, the following tasks were accomplished: 1) the facility was converted to using liquid N_2O as the oxidizer and Plexiglas® or HTPB as the fuel, which allowed for significantly higher nozzle pressure ratios and higher thrust; 2) the hybrid rocket motor chamber was redesigned to allow either a single hollow cylindrical fuel grain or an annular combustion volume formed by a smaller solid cylindrical fuel grain inside a larger hollow cylindrical fuel grain [see Figs. 1 through 4]; 3) the redesigned hybrid rocket motor was installed on an existing six-component thrust measurement stand similar to the one used in our cold-flow blow-down tests [see Fig. 5], and 4) twenty-two hot fire tests of this hybrid rocket motor were conducted [see Fig. 6 for typical first-year results]. In testing the single grain configuration at this point, we were having problems with our thrust measurement procedure, and we were

[1] Autodecomposition of N_2O, possibly exacerbated by the scale of the operation and by the high summer temperatures at Mojave, is the most likely factor in this accident. At the temperatures in our laboratory and at the scale of our experiments, autodecompostion of N_2O is self-quenching and therefore not a problem (See Rhodes, G.W., "Investigation of Decomposition Characteristics of Gaseous and Liquid Nitrous Oxide," Report AD-784 802, Air Force Weapons Laboratory, Kirtland Air Force Base, New Mexico, July 1974).

[2] "Optimal Thrust Vectoring for an Annular Aerospike Nozzle," a NASA Small Business Technology Transfer (STTR) project in which Cal Poly has teamed up with Rolling Hills Research Corporation of El Segundo, CA. Phase I was funded for $100,000 in January 2005 and completed in January 2006; Phase II was funded for $614,000 in June 2006 and will be completed in June 2008; a concurrent Phase III supplement for $39,942 was funded in August 2006, and a second concurrent Phase III supplement for $40,000 was funded in December 2007.

having difficulty preventing significant erosion of the nozzle throat in cases where the chamber pressure was over 200 psia. From these preliminary results, our conclusion was that high oxygen content in the exhaust gases, as indicated by the mixture ratio, was responsible for high nozzle throat erosion rates when the chamber pressure was over 200 psia.

In the second year of this project, after a complete reassessment of the safety of our testing procedure, we continued our iterative development process. We addressed the problems encountered in the first year of the project, and we explored ways of scaling up our hybrid rocket motor design to achieve higher thrust. In continuation of the project from year one, another fifteen hot fire tests of this hybrid rocket motor were conducted, and we accomplished the following goals: 1) the thrust measurement issue was solved; 2) we achieved average chamber pressures of nominally 300 psia with the hollow cylindrical Plexiglas® fuel grain, with improved nozzle throat erosion [see Figs. 7a]; and 3) achieved average chamber pressures of nominally 300 psia with the hollow cylindrical HTPB fuel grain, with even better throat erosion than with Plexiglas® [see Fig. 7b].

To pursue higher thrust, we tried two approaches to doubling the fuel grain area: 1) extending the length of the combustion chamber so that we could burn double-length fuel grains, and 2) developing the annular combustion chamber hybrid rocket motor [see Fig. 8]. The hot fire test results for the double-length Plexiglas® grain are shown in Fig. 9a. The double length provided twice the grain area, making the mixture ratio on the fuel-rich side of stoichiometric. Furthermore, the double length allowed for more complete combustion. Together these factors reduced nozzle throat erosion to a negligible level, 0.0003 in/sec on the nozzle throat radius. With throat erosion under control, the burn was very well behaved, with both chamber pressure and thrust at high levels and essentially constant.

The hot fire test results for the annular configuration are shown in Fig. 9b and the burned fuel grain is shown in Fig 10. Although these results are good, the mixture ratio was slightly oxidizer rich, and erosion of the nozzle throat was evident. Nonetheless, the annular configuration, which will be necessary to provide thermal insulation for the retaining rod holding an aerospike nozzle, shows promise. Testing of double-length or longer annular configurations using Plexiglas® and HTPB are planned and should show improved performance and negligible nozzle throat erosion.

Photos of typical hot fire tests of our hybrid rocket motor are shown in Fig. 11.

This project has been a complete success. Not only will this facility be beneficial for the testing of rocket nozzles, but the marriage of hybrid rocket motors and thrust vectored nozzles is interesting in its own right. Building on the success of this project and the enthusiasm it has generated in the undergraduate students, the graduate students, and the faculty, we are planning to design and build flight-weight hybrid rocket motors and composite airframes so that we can flight test our motors. As this effort builds, students and faculty from several departments in the Cal Poly College of Engineering will be involved. In the future, hybrid rocket motors equipped with advanced nozzles with thrust vectoring and throttling capabilities could play a very important role in programs trying to achieve inexpensive access to low earth orbit for NASA and the commercial sector, as well as becoming an inexpensive alternative for tactical missiles.

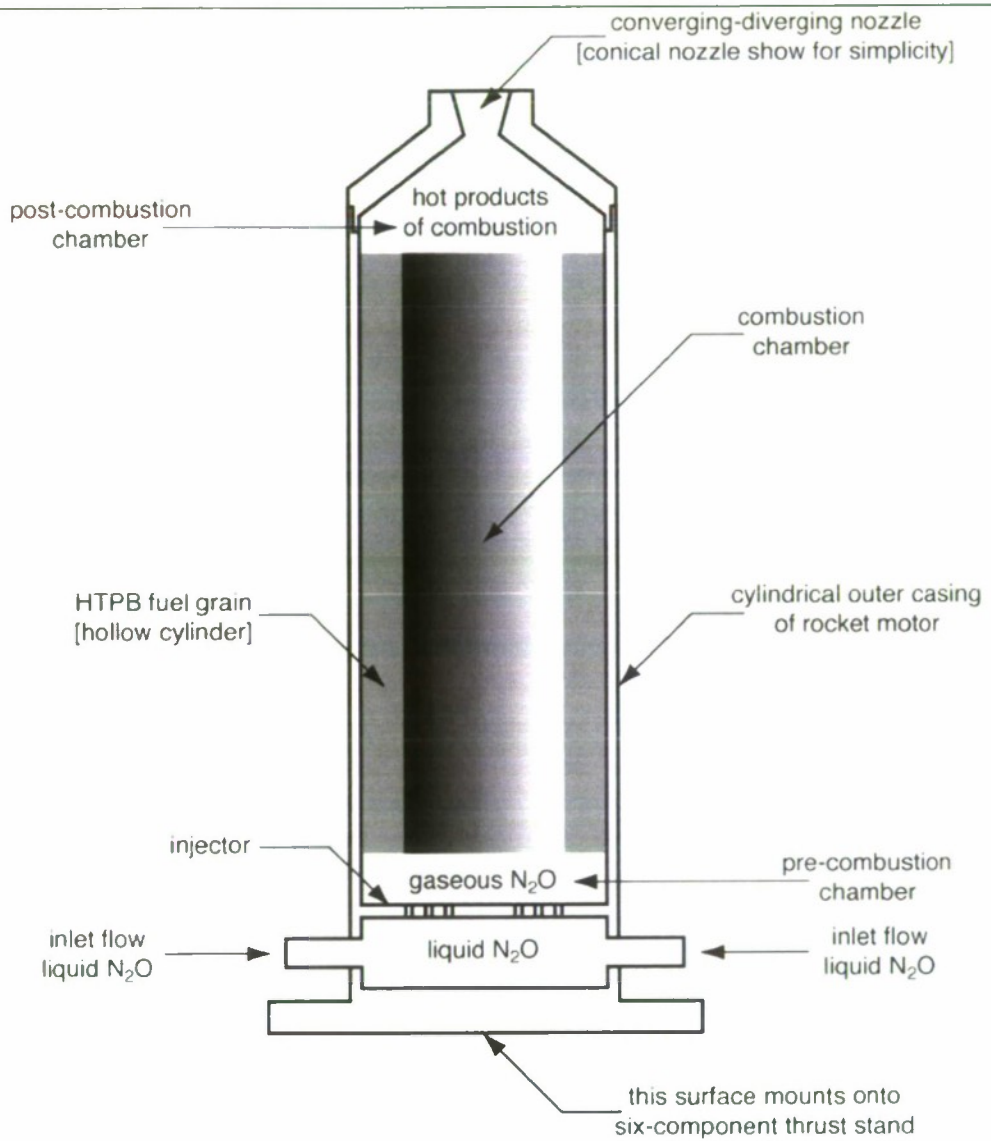


Figure 1: Schematic diagram of our hybrid rocket motor, shown with a conical nozzle and a single hollow cylindrical fuel grain.

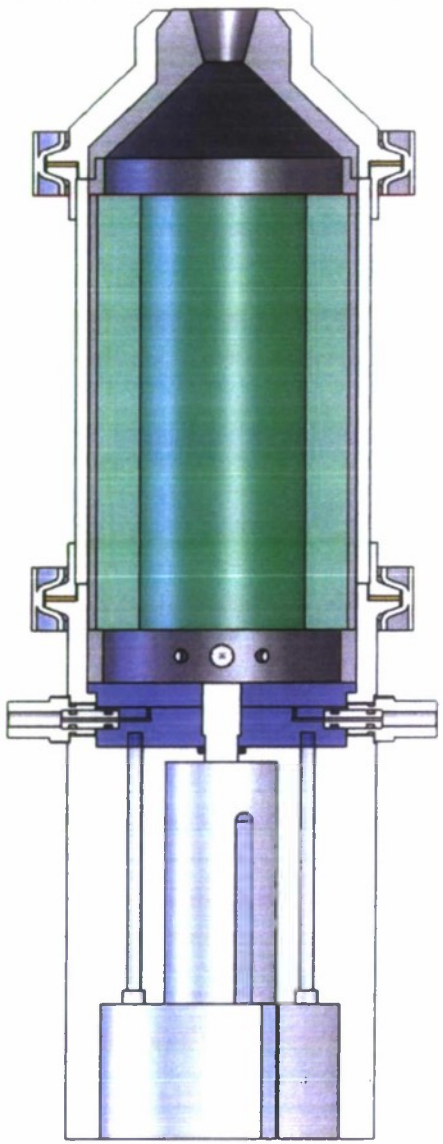


Figure 2. Design rendering of our hybrid rocket motor with a conical nozzle and a single hollow fuel grain.

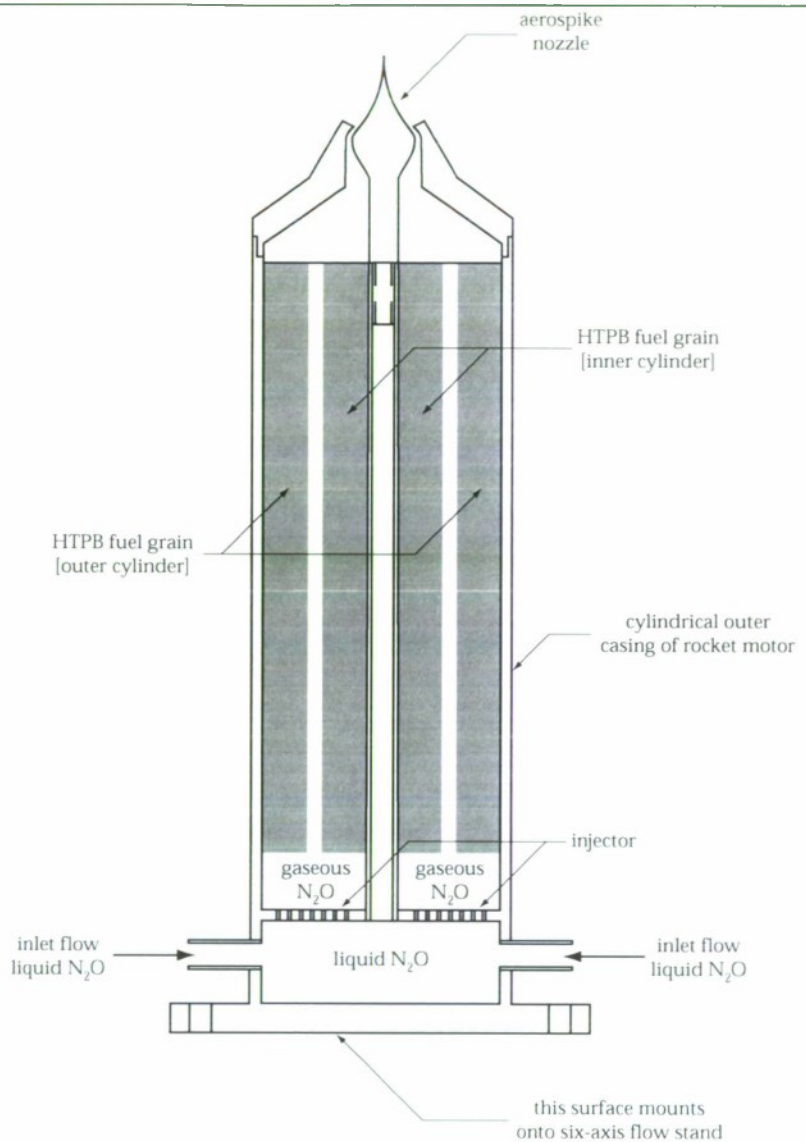


Figure 3: Schematic diagram of our hybrid rocket motor, shown with an annular aerospike nozzle and an annular combustion chamber. The inner grain provides thermal insulation for the rod that holds the aerospike in place.

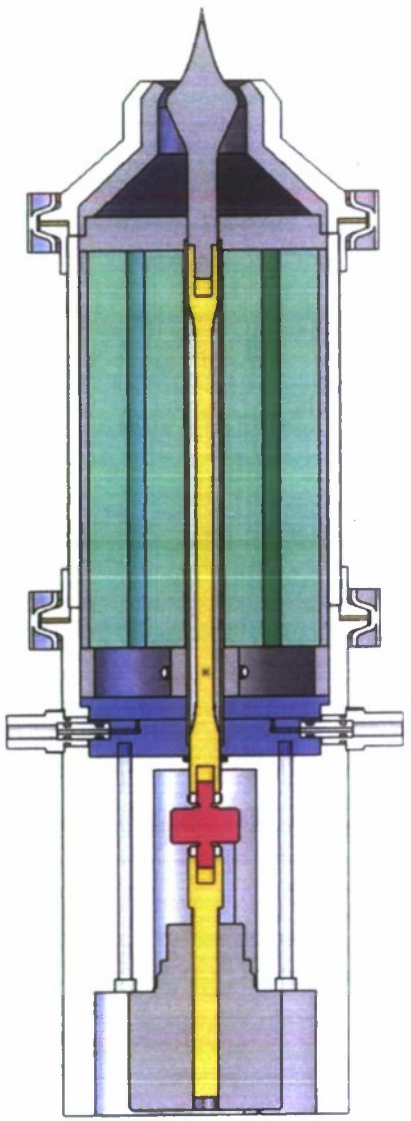


Figure 4. Design rendering of our hybrid rocket motor with an aerospike nozzle and an annular combustion chamber created between two cylindrical fuel grains.

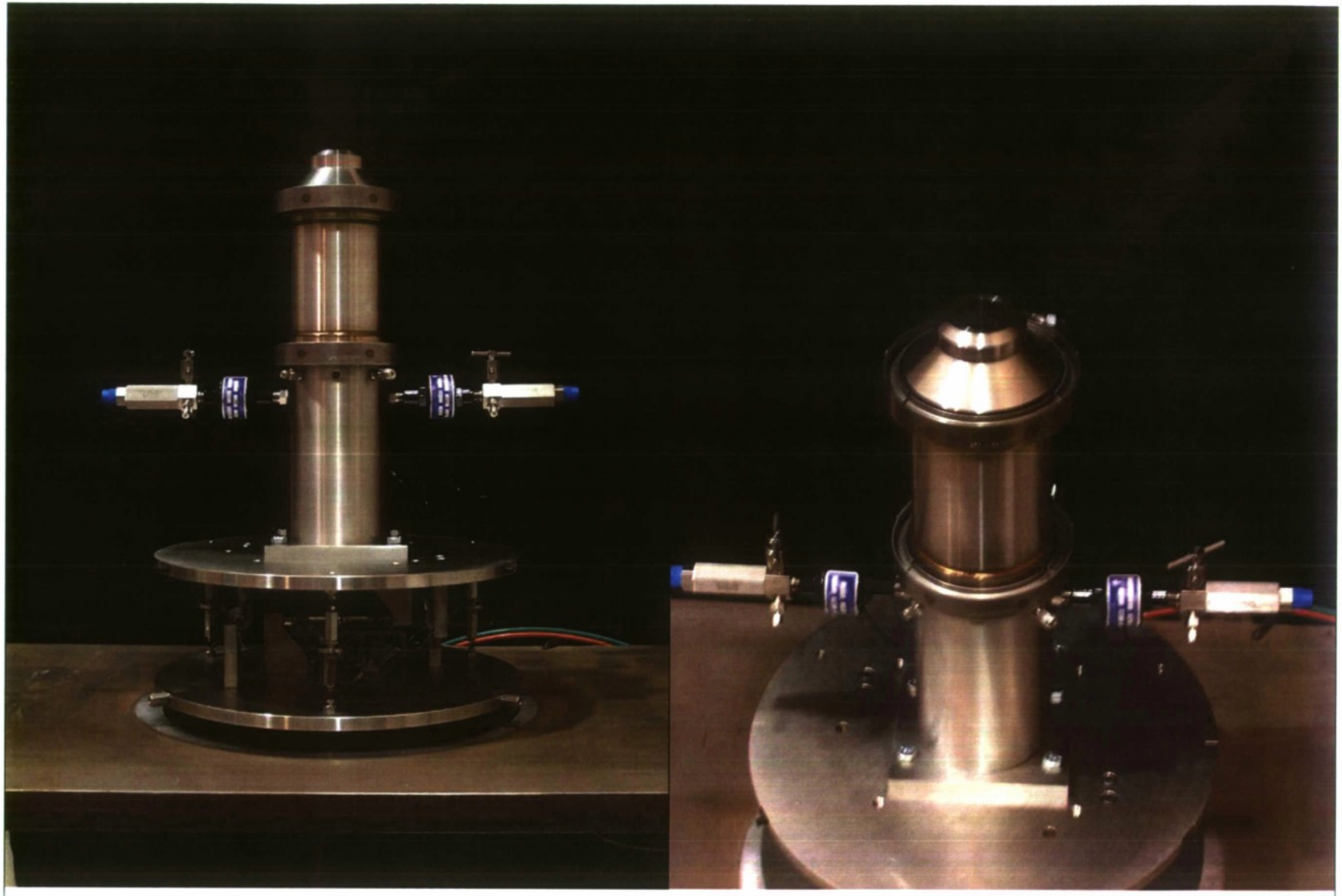


Figure 5. Two views of our hybrid rocket motor mounted on a six-component thrust stand. The outer diameter of the rocket motor housing is 4 inches.

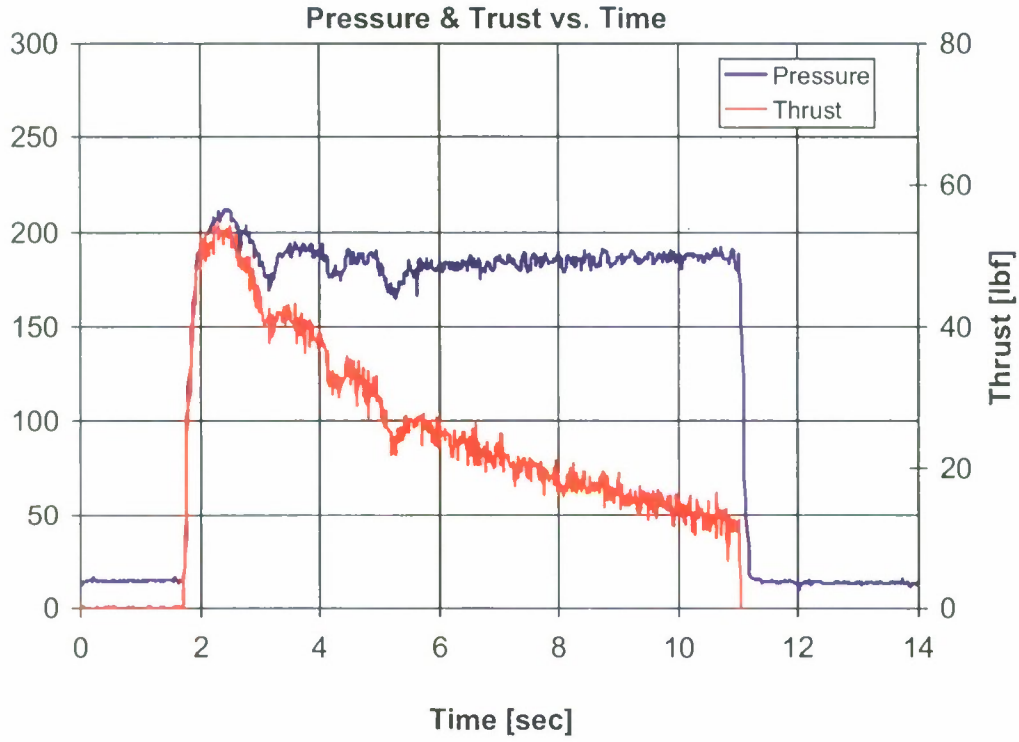


Fig. 6a. Mixture ratio of 4.17, which is somewhat fuel-rich relative to the stoichiometric mixture ratio of 5.28. Nozzle throat erosion: imperceptible over a 9-second burn.

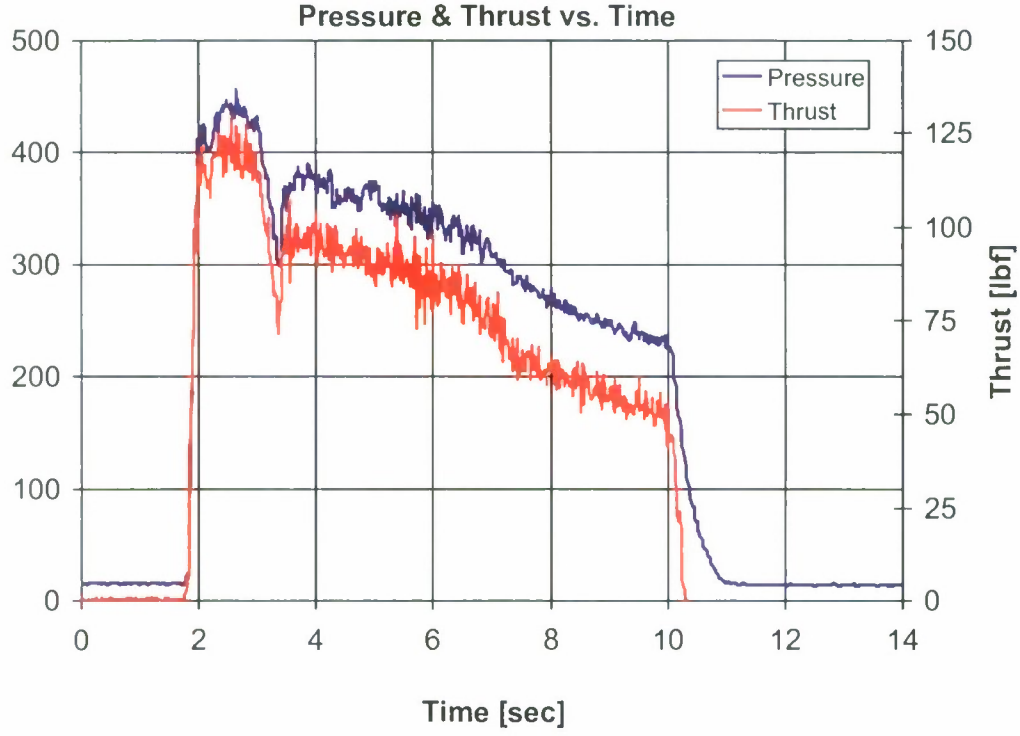


Fig. 6b. Mixture ratio of 8.66, which is significantly oxidizer-rich relative to the stoichiometric mixture ratio of 5.28. Nozzle throat erosion: Δradius = 0.00390 in/sec; Δarea = 25.21% total for an 8-second burn.

Figure 6. Chamber pressure and thrust versus time for hot-fire tests using liquid N_2O as the oxidizer and a hollow Plexiglas® grain as the fuel.

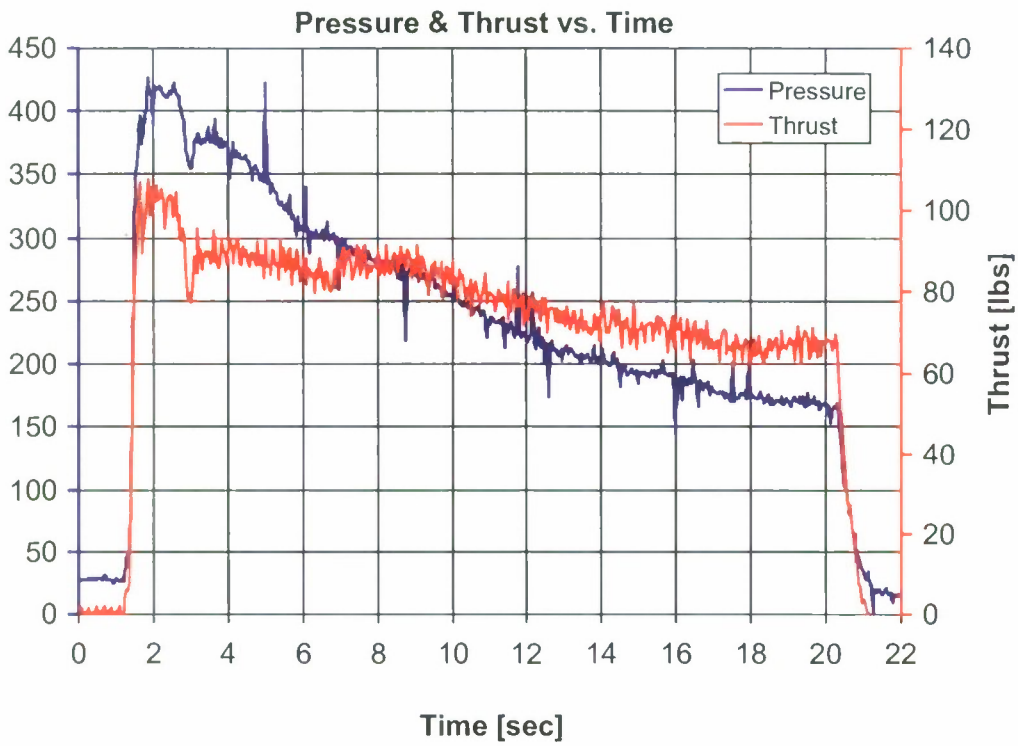


Fig. 7a. Mixture ratio of 7.87, which is significantly above the stoichiometric mixture ratio of 5.28. Nozzle throat erosion: Δradius = 0.00340 in/sec; Δarea = 55.26% total for a 19-second burn.

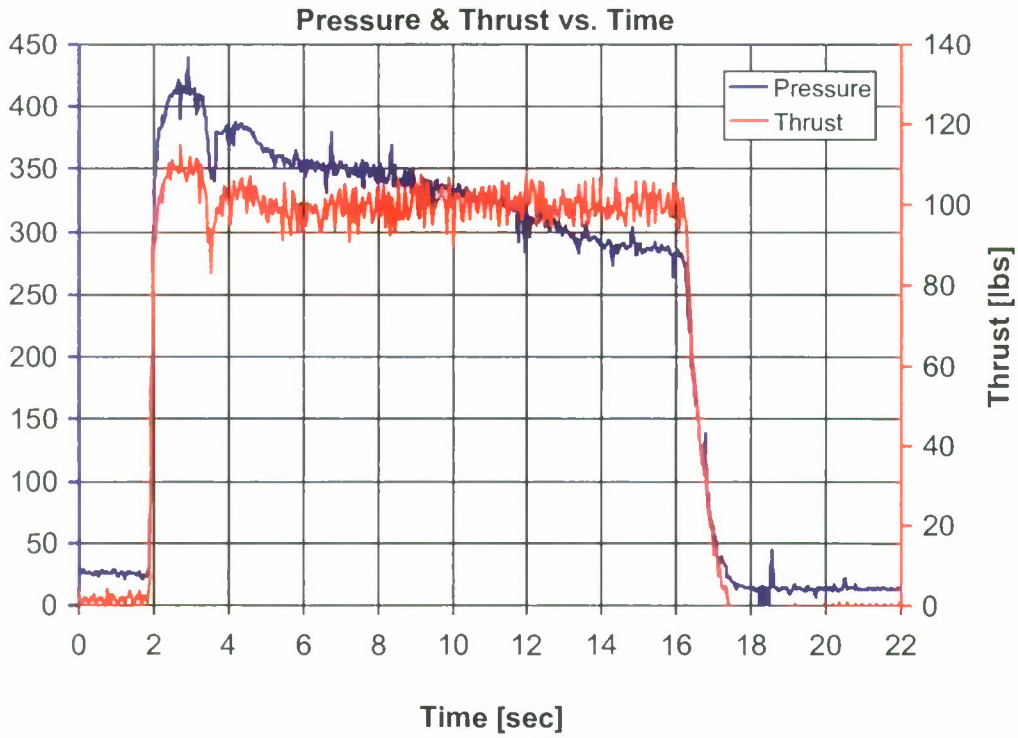


Fig. 7b. Mixture ratio of 7.87, which is somewhat below the stoichiometric mixture ratio of 8.85. Nozzle throat erosion: Δradius = 0.00277 in/sec; Δarea = 32.81% total for a 14-second burn.

Figure 7. Chamber pressure and thrust versus time for hot-fire tests using liquid N_2O as the oxidizer and a hollow Plexiglas® grain in Fig. 7a and a hollow HTPB grain in Fig. 7b.

Fig. 8a. Inner and outer Plexiglas® fuel grains that will form an annular combustion chamber.

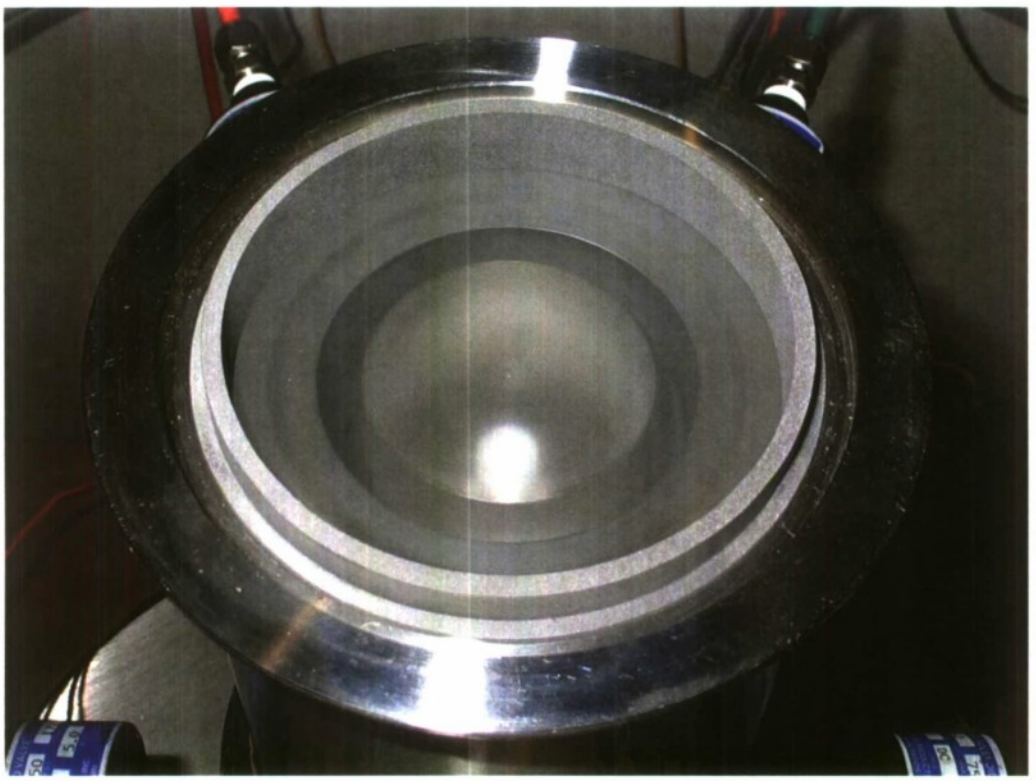


Fig. 8b. Dual Plexiglas®grain configuration shown installed in our hybrid rocket motor.

Figure 8. Our hybrid rocket motor configured for an annular combustion chamber that is formed by two Plexiglas® fuel grains: an inner solid cylindrical grain and an outer hollow cylindrical grain.

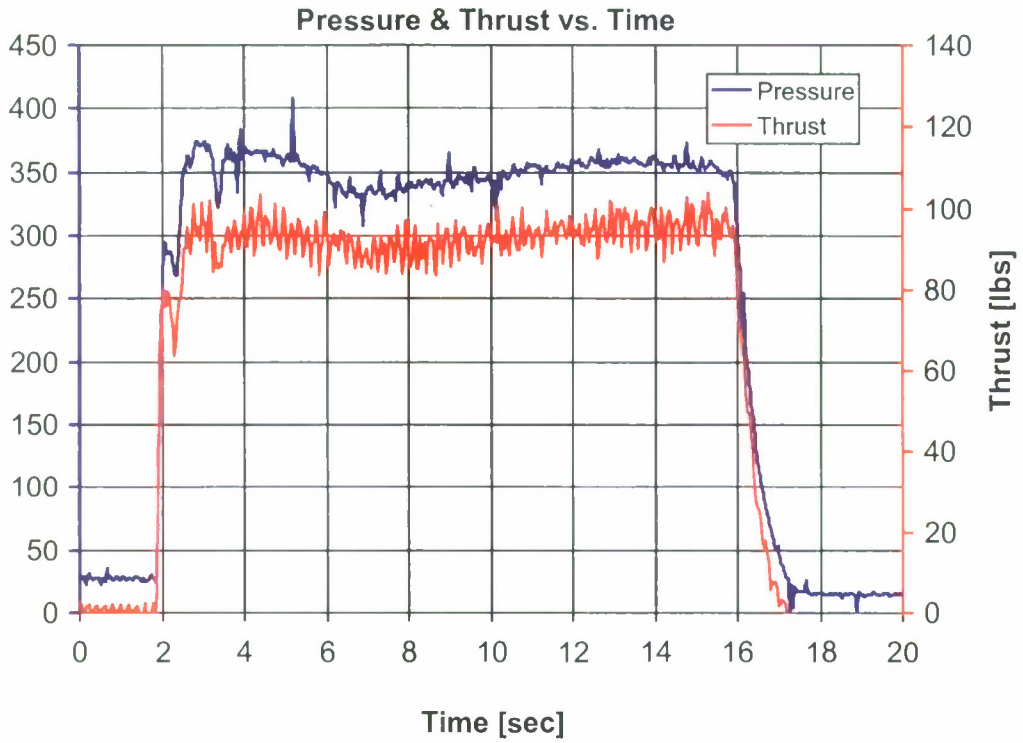


Fig. 9a. Mixture ratio of 4.63, which is somewhat below the stoichiometric mixture ratio of 5.28. Nozzle throat erosion: Δradius = 0.00023 in/sec; Δarea = 0.64% total for a 14-second burn.

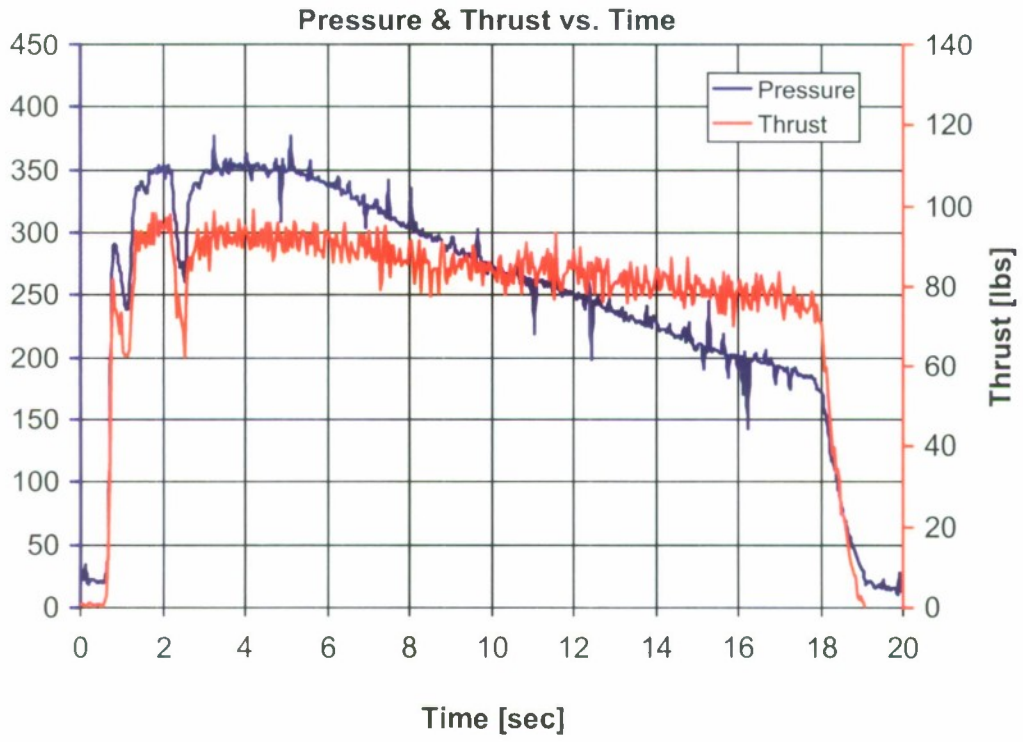


Fig. 9b. Mixture ratio of 5.96, which is slightly above the stoichiometric mixture ratio of 5.28. Nozzle throat erosion: Δradius = 0.00520 in/sec; Δarea = 79.44% total for a 17-second burn.

Figure 9. Chamber pressure and thrust versus time for hot-fire tests using liquid N_2O as the oxidizer and a double-length hollow Plexiglas® grain in Fig. 9a and pair of Plexiglas® grains forming an annular combustion chamber in Fig. 9b.

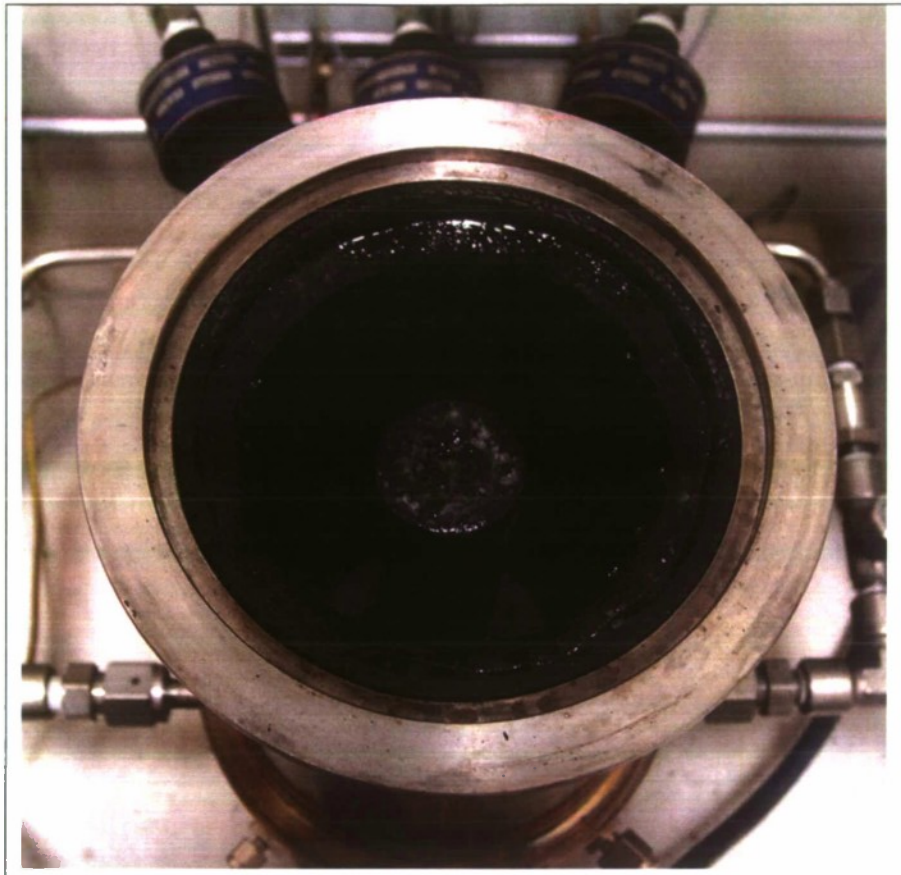


Fig. 10a: Inner and outer Plexiglas® fuel grains.

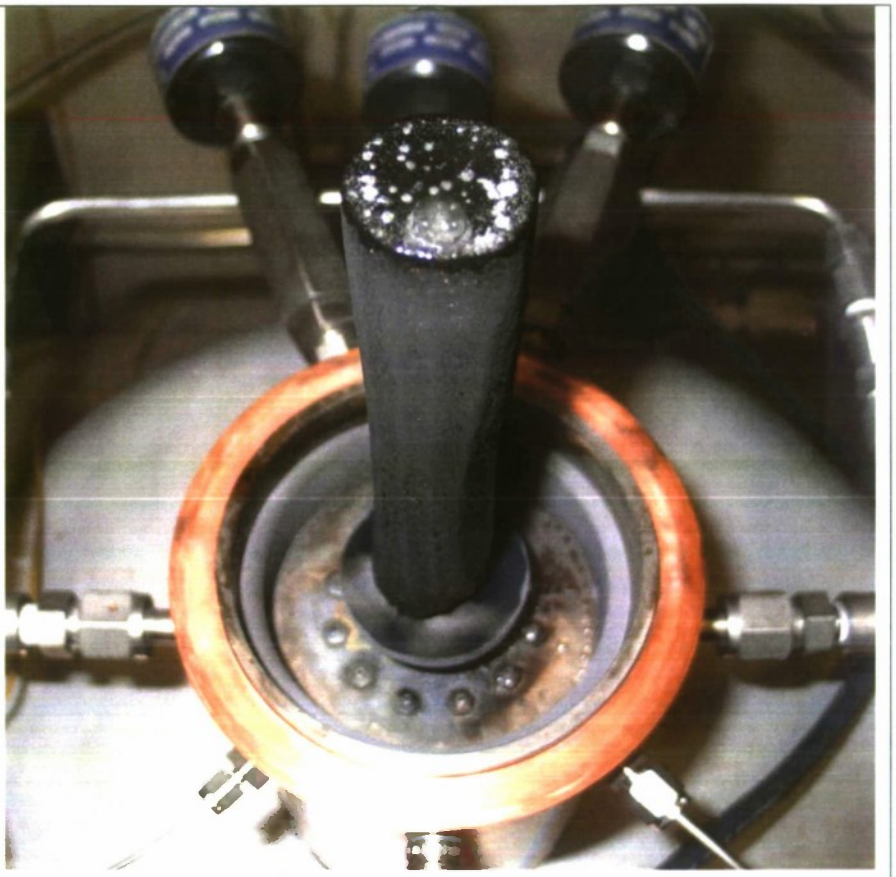


Fig. 10b: Inner Plexiglas® fuel grains.

Figure 10: Plexiglas® fuel grains after test firing. In total, 7.0 lbs. of N_2O and 1.17 lbs. of fuel grain were consumed in the 17.2 second burn. The fuel regression rates were 0.022 in/sec, 0.025 in/sec, and 0.029 in/sec for the outer radius, inner radius and inner length, respectively. Compare to the unburned fuel grains shown in Fig. 8.

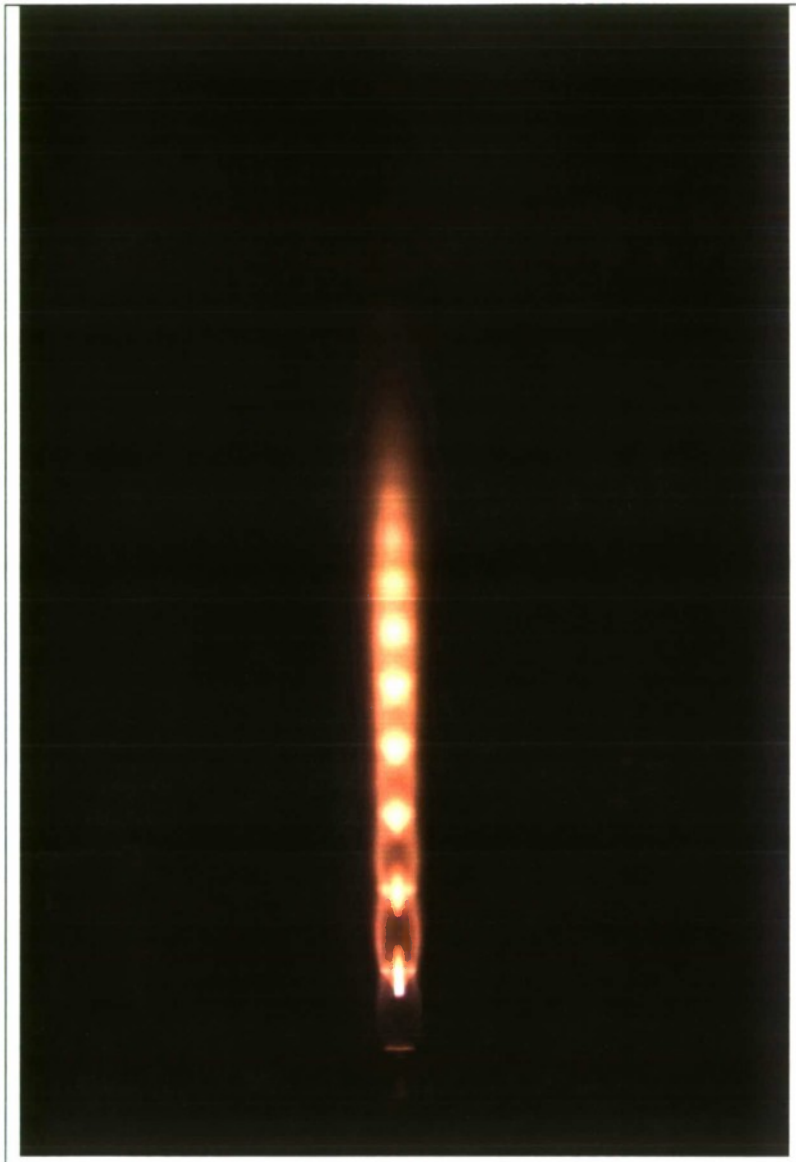


Fig. 11a: Firing of our hybrid rocket motor with N_2O and Plexiglas®.

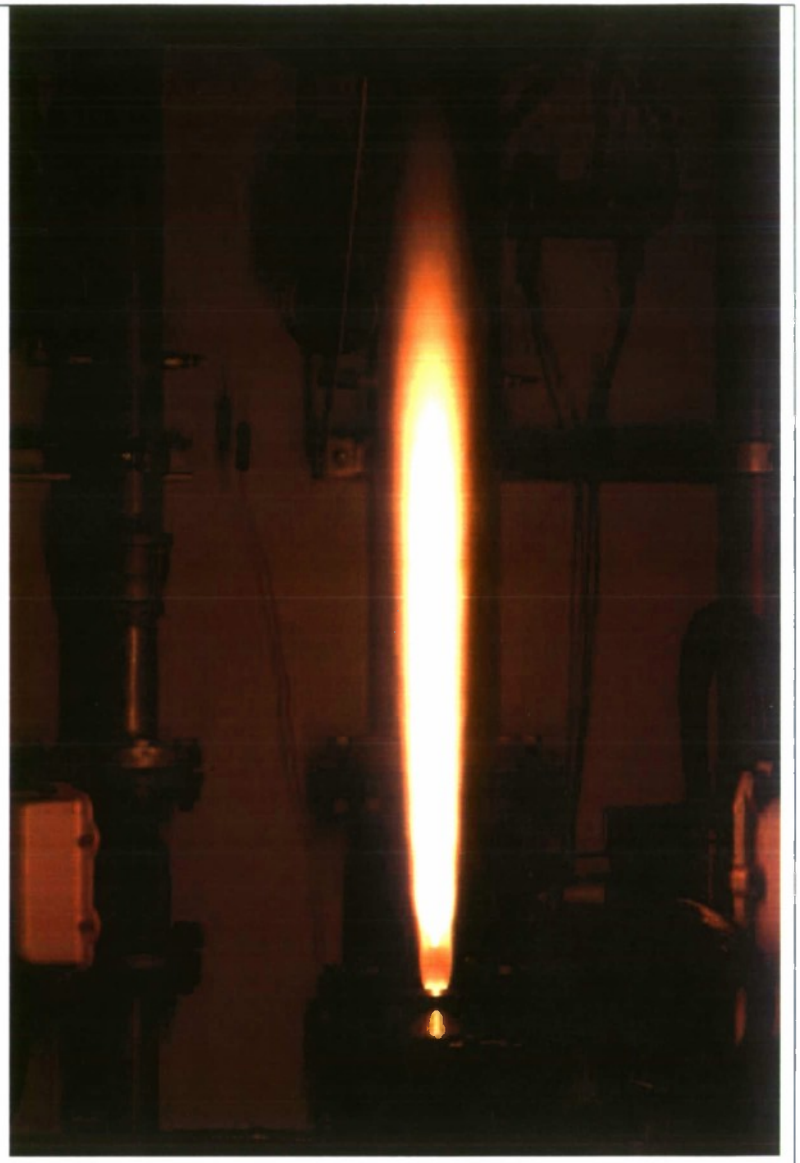


Fig. 11b: Firing of our hybrid rocket motor with N_2O and HTPB.

Figure 11: High-speed photographs of the hybrid rocket motor exhaust plume. The shutter speed was faster for Fig. 11a than for Fig. 11b.

Recycled Waste Materials as Additives to Improve the Performance of Cement-Treated Bases

Project Investigator:

Ashraf M. Rahim
Civil and Environmental Engineering
California Polytechnic State University
San Luis Obispo, CA

Abstract

Research was conducted to investigate the performance of various cement-treated aggregate mixtures mixed with industrial byproducts additives. Criteria based on the 7 and 28-day compressive strengths, were adopted on which each optimal supplemental material percentage was selected (350-450 psi @ 7 days). A series of unconfined compressive strength tests were performed along with splitting tensile strength tests. A threshold was not set for the splitting tensile test results, but this parameter was determined because tensile strength could help to restrain shrinkage. A modified version of ASTM D 559-03 (wetting and drying durability) was performed on supplemental material combinations which met the strength criteria.

Four industrial byproducts were evaluated in this research– Class F fly ash, Class C fly ash, Cement Kiln Dust (CKD), and Ground Granulated Blast Furnace Slag (GGBFS). It was found that the four different byproduct additives could partially replace cement in mixtures while still meeting the strength criteria. With the exception of fly ash class F all other byproduct additives enhanced the tensile strength. Mixes with fly ash class F could help achieving a desirable crack pattern. After twelve wetting and drying cycles mixes with fly ash class C lost the least strength followed by mixes with fly ash class F. Mixes with GBFS suffered the greatest loss in strength after wetting and drying cycles.

ACKNOWLEDGMENT

This report includes the results of a study titled, "Recycled Waste Materials as Additives to Improve the Performance of Soil-cement – A Laboratory Investigation" sponsored by the Department of the Navy, Office of Naval Research, under Award # N00014-05-1-0855.

The author wishes to thank Dr. Susan Opava for making this support available. The hard work of under-graduate students Alex Ubaldo and Bill Johnson who worked on this project and input from Jennifer Beck in preparing this report are also acknowledged. The author is greatly thankful to Mr. Kirk McDonald of Colton Cement Plant for donating the CKD and his help in obtaining the fly ash used in this study. Mineral Resource Technologies, Inc. donated fly ash class C and class F.

INTRODUCTION

The American Concrete Institute defines soil-cement as a mixture of soil and measured amounts of portland eement and water eompaeted to a high density (ACI 1990). This definition can be further broadened to inelude other aggregates besides soil. As in concrete, aggregate particles are bonded together by the cement paste. However, unlike the ease of eonerete individual partieles are not eompletely eovered by the cement paste.

Cement-treated base (CTB) is a eompacted mixture of graded aggregate, eement and water. This mixture is used as abase eourse for the eonstruction of highways, airport runways and taxiways. A CTB is used to provide a strueturally sound base material over weak sub-grades such as expansive clay (http://gearsinc.com/prod02.htm, aecessed 12/30/2007). Cement-treated aggregate is appealing as a low-cost construction material, and investigations have examined the use of marginal aggregates in combination with sand and cement, for use in low-cost pavement bases (Majumder et al. 1999). Other researchers have also studied the use of reeyeled eonerete in CTA (Lim and Zollinger 2003).

Most problems with cement-stabilized base layers in pavement stem from the fact that eurrent design praetiees are based only on strength (Guthrie et al. 2006). For example, many state department of transportation (ineluding Caltrans) require sufficient cement to achieve a minimum uneonfined compressive strength as high as 750 psi after seven days. While this level of eement results in a very stiff base layer, high dry shrinkage induced stresses could lead to cracking of the base layer. These cracks within heavily cement-stabilized base layers refleet into the surfaee asphalt concrete layer as transverse craeks.

As the cost of cement began to rise in the 1970s other additives/byproduct materials that possess cementitous properties are tried by partially replacing the cement. Examples of these additives/byproduct materials include fly ash, cement kiln dust, granulated blast furnace slag among others.

OBJECTIVE

Reducing the potential for cement treated bases to develop shrinkage cracks is crucial in order to improve the pavement serviceability and for a cost effective design. Recent studies suggest that crack-related degradation can be abated by adopting materials and/or methods that bring about a "desirable" crack pattern, "desirable" being defined as numerous fine cracks at close spacing, which ensures adequate load transfer across the cracks (George 2002). It is not so much the number of cracks but the width of these cracks that has a significant influence on the long-term performance of the pavement since wider cracks have the tendency to reflect through the overlying pavement. Limiting/controlling drying shrinkage can effect the development of this "desirable" crack pattern in a stabilized layer. Several alternatives are available to control the drying shrinkage. These include: judiciously selecting the cement dosage, selecting a soil for stabilization having a limited fines content and plasticity, and the use of additives in conjunction with Portland cement, all of which promote development of a "desirable" crack pattern in a stabilized layer. The main objective of this study is to investigate the effect of four byproduct additives on the strength characteristics of cement-treated bases.

EXPERIMENTAL WORK AND RESULTS

Materials

Coarse, medium and fine aggregates were blended to produce a mix that meets Caltrans' gradation for CTB Class A (see Table 1). Portland cement type II was used in preparing all specimens tested in this study.

Table 1. Particle size distribution for aggregate used in the study

Sieve Size	Specification, %	% Passing
3-in	----	100
2 ½ -in	-----	100
1-in	100	100
¾-in	90-100	94
No. 4	40-70	68
No. 30	12-40	33
No. 200	3-15	5

Four industrial byproducts were included in this research— fly ash class F, fly ash class C, Cement Kiln Dust (CKD) and Ground Granulated Blast Furnace Slag (GGBFS). Historically, class C fly ash has been the preferred supplemental material for concrete because of its self-cementitious properties, but class F fly ash is produced in larger quantities and finds fewer commercial applications. The source of fly ash is coal-fired electric power plants where fly ash is produced as a combustion byproduct formed as flue gasses cool. According to the American Coal Ash Association, as of 2001, 58% of the coal combustion products were Class F fly ash and 68% of these required disposal for lack of use (Arora and Aydilek 2005). Fly ash class C and F used in this research were donated by Mineral Resource Technologies, Inc. Table 2 presents the chemical composition of the fly ash used in this study.

Table 2. Chemical composition of fly ash

	SiO2	Al2O3	Fe2O3	SO3	CaO	MgO
FA class C	36.10	17.98	4.66	2.37	23.33	3.96
FA class F	47.87	21.30	5.08	1.24	13.83	2.84

Cement Kiln Dust (CKD) is a byproduct of the cement production process. CKD is captured by air pollution eontrol equipment installed on the chimneys of cement kilns. Aeeording to the EPA in 1990, the cement industry generated an estimated 12.7 million metric tons of CKD at 111 plants in 38 states of which 4 million metric tons were disposed of in piles, quarries, and landfills. The industry disposed of an estimated 3.3 million metrie tons in 1995. There are eurrently 110 Portland eement plants operating in thc United States and Puerto Rico with the highest production in California, Texas, Pennsylvania, and Michigan. CKD has cementitious properties which are dependent on the eoneentration of hydratablc oxides such as free (unreacted) lime or free Magnesia (MgO). But its finc powdery nature (0.3 mm max. size) and caustic properties (pH≈ 12) make CKD difficult material to handle in bulk. If moisture is added to control dust some of thc cementitious properties are saerifieed (premature hydration) and the material becomes clumpy [Turner-Fairbanks Highway Research Center (TFHRC) 2002a]. It was included in the scope of this research because it is currently underutilized and has shown potential as a partial replacement for cement in mixturcs (Udoeye and Hyec 2002). CKD for this study was donated by Colton Cement Plant, Colton, California. The ehemical composition of CKD is shown in Table 3

Table 3. Chemical composition of CKD

Component	SiO2	AL2O3	Fe2O3	CaO	MgO	SO3	Na2O	K2O	Ti2O5	P2O5	LOI
%	12.86	4.46	0.91	51.11	2.39	2.52	0.40	1.72	0.07	0.20	22.20

Ground Granulated Blast Furnace Slag (GGBFS) is a by product of the steel industry. Blast furnace slag is defined as "the non-metallic product consisting essentially of calcium silicates and other bases that is developed in a molten condition simultaneously with iron". In the production of iron, blast furnaces are loaded with iron ore, fluxing agents, and coke. When the iron ore, which is made up of iron oxides, silica, and alumina comes together with the fluxing agents, molten slag and iron are produced. The molten slag then goes through a particular process depending on what type of slag it will become. Air-cooled slag has a rough finish and larger surface area when compared to aggregates of that volume which allows it to bind well with Portland cements as well as asphalt mixtures. GGBFS is produced when molten slag is quenched rapidly using water jets, which produces a granular glassy aggregate (Cervantes and Roesler, 2007).

Testing

The optimal moisture content (OMC) required for compaction of the cement-aggregate mixture was determined based on the optimum Unconfined Compressive Strength (UCS) of 4.0-inch diameter by 4.6-inch height cement-aggregate specimens prepared at 5% cement, compacted employing standard proctor compaction and cured for 3 days. The specimen's proportions were not the ideal height-to-diameter ratio (h/d) of 2.00 used for Method B of ASTM 1633-00— *Standard Test Methods for Compressive Strength of Molded Soil-Cement Cylinders*. When the h/d ratio is 2.0 the compressive strength values are more accurate because complex stress conditions present at lower h/d values are reduced. This is, however, the standard mold size used for both the moisture-density and the wetting and drying durability tests. Test Method A of ASTM 1633 uses this common mold size which has a h/d ratio of 1.15. For comparison with samples molded with the

ideal shape, a correction factor based on ASTM C 42 for conerete would use a strength reduetion of 1/(1.1) (ACI 1990). As shown in Figure 1 the optimum moisture content (OMC) is approximately 6%.

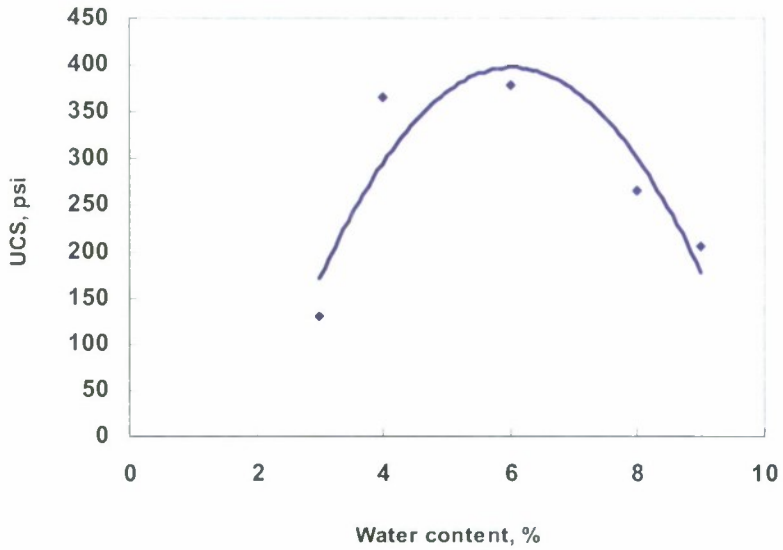


Figure 1. Water content vs. UCS for 5%-cement mixture

For eement-aggregate mixtures with byproduet additives a portion of the 5% cement was replaeed by an estimated equivalent portion of the byproduet additives. Several mixes with byproduct additives were tested for UCS and the mixtures that met the strength criterion were seleeted. This test was performed in general aeeordanee with ASTM D 1633-00, a compressive strength test for molded eement-aggregate eylinders. Specimens from the same initial set of eylinders were also tested for their split tensile strength in aeeordanee with ASTM C 496-04 designed for eonerete. Speeimens were eured at 95% humidity room for 3, 7 and 28 days before being tested. Table 4 lists the suceessful mixtures ineluded in the study.

Table 4. Successful mixes with cement and byproduct proportions

Designation	Description	Cement, %	Byproduct, %	Water content, %
C	Cement-only mix	5	------	6.0
C/CKD	Mixes with cement and CKD byproduct	2	15	7.5
C/FAC	Mixes with cement and fly ash class C	4	6	6.5
C/FAF	Mixes with cement and fly ash class F	2	12	6.5
C/GGBFS	Mixes with cement and GGBFS	5	0.5	6.0

Unconfined Compressive Strength (UCS)

Figures 2 presents the UCS test results. As noticed from Figure 2 all mixes with byproduct additives met the 7-day strength criterion adopted in this study (350-500) psi. By replacing 4.5% of the cement (expensive material) needed in the cement-only mix by 5% GGBFS (much cheaper byproduct) the aggregate mixture achieved approximately the same 7-day UCS. Mixes with fly ash class F exhibited the slowest strength gain trend among all mixes tested in this study. This trend could help in achieving the desirable crack pattern (several closely spaced minute cracks taking place early after construction) minimizing the possibility of reflection through the asphalt surface layer.

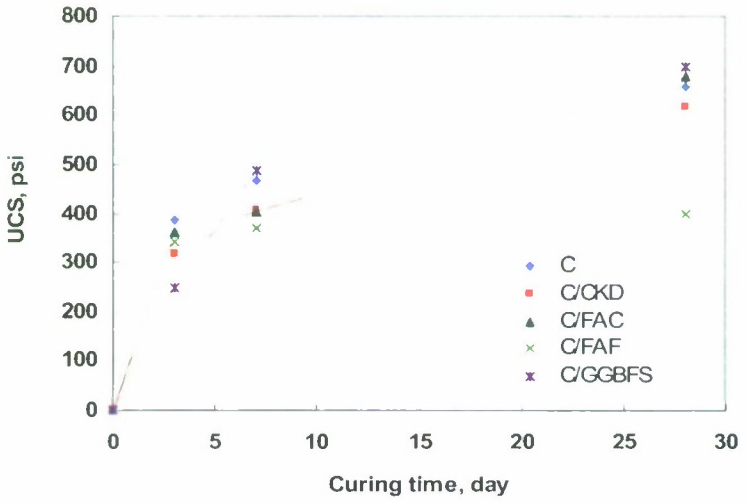


Figure 2. UCS for different mixes at different curing periods

Indirect Tensile Strength (ITS)

The tensile strengths of seleeted mixture designs were evaluated using ASTM C 496-04—*Standard Test Method for Splitting Tensile Strength of Cylindrical Concrete Specimens*, a test designed for eonerete whieh measures the tensile performance of a speeimen indireetly.

In the test setup shown in Figure 3 the eement-aggregate cylinder is held on its side sandwiehed in a jig. This produees loading whieh induees tensile stresses in the plane of applied loading. The test maehine used for this proeedure was also used for the unconfined eompressive strength testing. Plywood bearing plates were added to distribute the load and prevent speeimen movement.

Figure 3. Machine used for ITS and UCS test

It is seen in Figure 4 that Mixes with fly ash elass F exhibited the lowest ITS gain during the first 2 weeks. This result suggests that drying shrinkage indueed tensile stresses eould overeome the tensile strength of this mix early after construetion produeing a desirable eraek distribution (elosely spaeed minute eraeks) rather than wide eraeks at larger spacing

which could reflect at a faster rate in the asphalt concrete surface. Narrow closely spaced cracks are less likely to translate through asphalt overlays built on cement-treated bases. Higher tensile strength allows greater stresses to develop before cracking is initiated. This leads to larger more widely spaced cracks which can reflect through pavement (George, 2002).

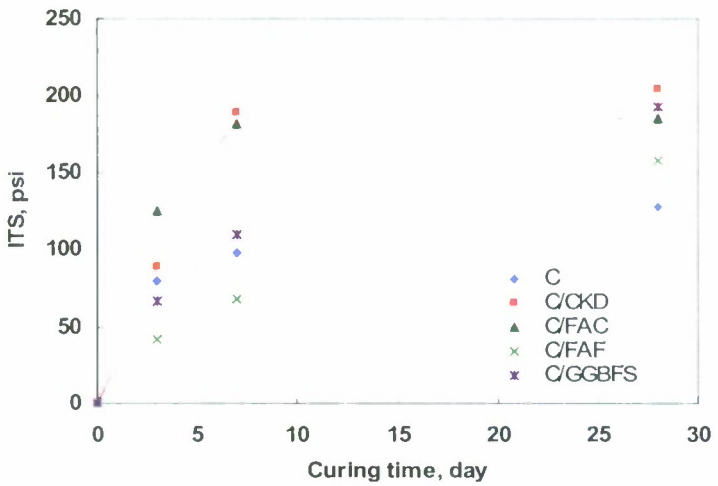


Figure 4. ITS for different mixes at different curing period

Wetting and Drying Durability

A modified ASTM D 559-03—*Standard Test Methods for Wetting and Drying Compacted Soil-Cement Mixtures* was performed to evaluate the durability of the different mixtures. In this this modified version sample no sample were subjected to the manually conducted heavy brush stroks which is considered subjective. However, the two samples were exposed to a series of 12 cycles ofalternating immersion in room temperature water for 5 hours followed by drying in an oven at 160°F for 42 hours. This is followed by UCS testing conducted on both samples. The UCSs determined from test were compared with the 28-day strength for same mixes without being exposed to wetting/drying. The loss in UCS was used as a measure of mixes durability (see Table 5).

Table 5. Loss in UCS after twelve wetting and drying cycles

Mix	Base UCS[a]	WD-UCS[b]	% Loss
C	656	523	20.0
C/CKD	618	351	43.0
C/FAC	455.	466	5.0
C/FAF	398	323	19.0
C/GGBFS	698	327	53.0

a 28-day UCS without wetting drying
b UCS after wetting drying

As seen in Table 5, mixes with fly ash class C experienced the lowest loss in UCS following wetting and drying cycles. Mixes with fly ash class F and cement-only ranked second and third, respectively. Mixes containing GGBFS additive exhibited the poorest performance with regard to strength loss.

CONCLUSIONS

1. All compressive strengths for mixes with byproduct additives met the 7-day strength criterion set for this study.

2. The addition of GGBFS additive allowed a significant reduction in the amount of cement required to meet the compressive strength criterion established for this test program.

3. Mixes with CKD and fly ash class C exhibited the fastest rate of increase in strength gain while mix with fly ash class F experienced the slowest strength gain making it a candidate for treated base that could provide a desirable crack pattern.

4. Based on the loss of UCS after wetting-drying durability test, Mix with fly ash class C was ranked the best, followed by mix with fly ash class F and mix with cement only.

5. More research is needed to study the effect of these additives on shrinkage characteristics. Field trials will be needed to verify the findings of this study in-situ.

REFERENCES

American Concrete Institute (ACI) (1990), *State-of-the-Art Report on Soil-Cement,* ACI 230.1R-90, Committee 230, Farmington Hills, MI.

Cement Treated Base (2007), (http://gearsinc.com/prod02.htm, accessed 12/30/2007).

Majumder, B. K., A. Das and B. B. Pandey (1999). Cement Treated Marginal Aggregates for Roads. *Journal of Materials in Civil Engineering*, 11(3), 257-265.

Lim, S. and D. G. Zollinger (2003). Estimation of the Compressive Strength and Modulus of Elasticity of Cement-Treated Aggregate Base Materials. *Transportation Research Record* (1837), 30.

Guthrie, W. S., Sebesta, S. and Scullion, T. (2006), Improving Long-term Performance of Cement-Treated Aggregate Base Materials. Texas Transportation Institute, *Report No. 7 4920-S.*

George, K.P. (2002), "Minimizing Cracking in Cement Treated Materials for Improved Performance," *Research and Development Bulletin No. RD123,* S Portland Cement Association Skokie, Ill

Arora, S. and Aydilek, A.H. (2005), "Class F Fly-Ash-Amended Soils as Highway Base Materials," *Journal of Materials in Civil Engineering,* Vol. 17, No. 6 November/December, pp. 640-649

Turner-Fairbanks Highway Research Center (TFHRC) (2002a), "Coal Fly Ash User Guideline—Stabilized Base," http://www.tfhrc.gov/hnr20/recycle/waste/cfa55.htm, Accessed on April 30, 2006

Udoeyo, F. F. and Hyee, A. (2002), "Strengths of Cement Kiln Dust Concrete", *Journal of Materials in Civil Engineering*, Volume 14, Issue 6, pp. 524-526

Cervantes, V. and Roesler, J. (2007), Ground Granulated Blast Furnace Slag, Technical Note No. 35 http://www.ceat.uiuc.edu/PUBLICATIONS/technotes, Accessed 12/30/2007

Enhanced Operator Control for Remotely Operated Vehicles

Project Investigator:

Louis Rosenberg
Graduate Studies in Education
California Polytechnic State University
San Luis Obispo, CA

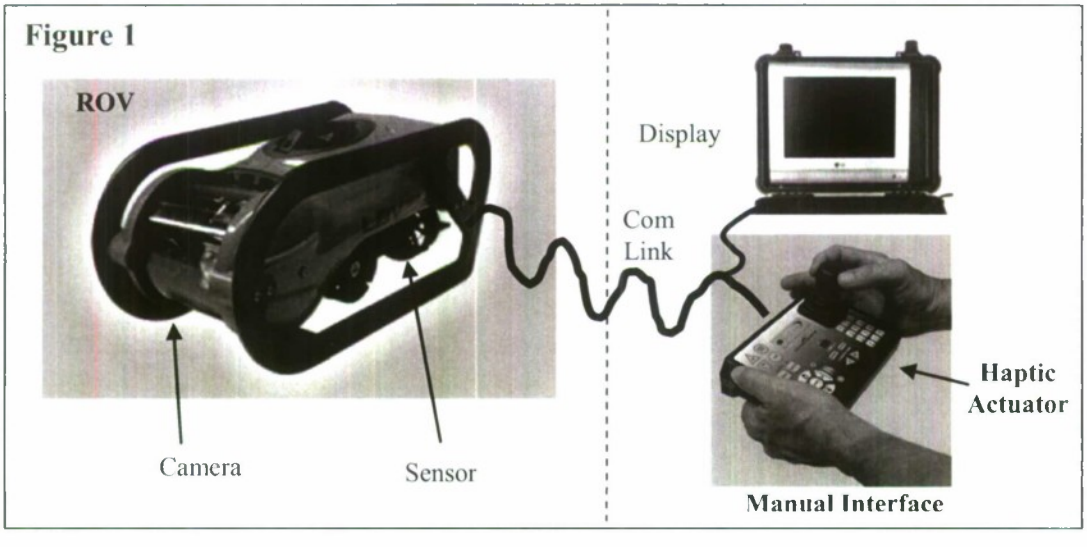
Project Objective.

The goal of this research project was to explore how tactile feedback can be used to improve user control of remotely operated vehicles (ROVs) in underwater environments. More specifically, this project has been focused on developing unique hardware, software, methods, and algorithms for enhancing a human operator's ability to guide an underwater vehicle in a turbulent offshore environment. The project has been split into three phases: Phase 0, Phase I, and Phase II. During the **Phase 0** effort, a remotely operated vehicle (a Seabotix LBV150 ROV) was tested off Cal Poly research pier in Avila California to evaluate operator control issues in turbulent conditions. Based upon this initial testing, a number of tactile feedback strategies were formulated with the objective of addressing specific operator difficulties. During the **Phase I** effort, a hardware and software system was developed to (a) generate tactile feedback sensations on a human operator of a Seabotix ROV hand-controller, (b) collect real-time acceleration data from the Seabotix ROV during operation, and (c) create a real-time control system in which tactile sensations can be automatically triggered by acceleration events detected on the ROV. During the **Phase II** effort the final system will be tested off the Cal Poly research pier to evaluate if the tactile feedback provides human operators with sensor cues that they find helpful during typical ROV tasks.

System Configuration:

The basic system architecture, as shown in **Figure 1** below, includes a Seabotix ROV connected by a communication tether to an operator control station. The control station includes a display screen for viewing the images from the ROV cameras and a manual interface operator control. As to be developed during the current research project, sensors will be mounted on the ROV for detecting underwater events and a tactile feedback actuator will be mounted upon the manual controller for providing electronically controlled tactile sensations to the operator:

Figure 1



Project Summary:

A team of undergraduate students (Andrew Gray, William Krehbiel, and Hidemi Nishibayashi) were hired to support the Phase 0 and Phase I efforts under the supervision of Dr Rosenberg and with ROV support from the Cal Poly Center for Costal Marine Sciences headed by Dr. Moline. The Phase 0 effort was completed during the spring of 2007, during which time a team of student tested a Seabotix LBV150 ROV in turbulent conditions off the Cal Poly Pier. The goal of the effort was to observe and document the challenges to operator control posed by turbulent conditions and then propose tactile feedback strategies for addressing operator difficulties. The Phase 0 effort revealed the following observations:

a) During turbulent conditions, sudden sea swells would regularly move the ROV to a new underwater location through abrupt and unexpected events. This left the human operator disoriented, for the normally predictable motion of the ROV was disrupted by a sudden and unexpected shift in position. The disorientation was made worse by the fact that in murky conditions, there were generally few visual cues by which the operator could judge the distance the ROV had moved, or even if it had moved at all. In fact, the operators reported that often the ROV would surge to a new location and they were not be sure if anything had happened. Thus if was postulated that operator performance would be enhanced if human operators could be made aware when sea a swell hit the ROV as well as be made aware as to the magnitude of the sea swell so they could anticipate how far the ROV had been moved.

b) During turbulent conditions, especially in murky water, the human operators reported having difficulty determining if the ROV had collided with an obstacle. This was particularly true if the collision occurred from below, behind, or to the side, for the visual feedback provided to the operator is primarily frontal. In general the obstacles that the ROV would hit included the sea floor and the pier pilings. Thus it was postulated that operator performance would be enhanced if users could be made aware when the ROV had impacted an obstruction.

Based on the observations made during the Phase 0 testing as described above, a set of control strategies were formulated for using tactile feedback to enhance operator control. These strategies were aimed at providing helpful and intuitive cues to human operators that would enhance the standard visual and audio feedback that is normally provided from the ROV. The strategies that were proposed were as follows:

1) **Tactile Feedback triggered by Underwater Swells** – it was postulated that a helpful form of tactile feedback could be created in which a tactile cue is provided in response to the detection of a sea swell hitting the ROV. More specifically, a tactile cue could be provided when an accelerometer on board the ROV detects that the vehicle has been hit by a swell and moved with greater than a threshold level of acceleration. The tactile cue would thus inform the operator of a significant event that they otherwise would be limited information about or sometimes be totally unaware of. In addition, the magnitude of the tactile sensation would be scaled to the magnitude of the detected swell acceleration. In this way the operator could determine based upon the relatively intensity of the tactile sensation (in either magnitude, frequency, or both), the relative

intensity of the sea swell impact. With such feedback the operator would not only know that the vehicle was abruptly jarred by an invisible sea swell, the operator would have a sense of the relative magnitude of the event.

2) **Tactile Feedback triggered by Underwater Collisions** – it was postulated that a helpful form of tactile feedback could be created in which a tactile cue is provided in response to the detection that ROV has hit an obstruction. More specifically, a tactile cue could be provided when an accelerometer on board the ROV detects that the vehicle has collided with a stationary obstruction resulting in an acceleration impulse greater than a threshold level. The tactile cue would thus inform the operator of a significant event that they otherwise would be limited information about or sometimes be totally unaware of, especially if the collision occurs from behind, the side, or below. In addition, the magnitude of the tactile sensation would be scaled to the magnitude of the detected collision acceleration. In this way the operator could determine based upon the relatively intensity of the tactile sensation (in either magnitude, frequency, or both), the relative intensity of the collision impact.

Based on the above objectives, a Phase I effort was pursued to design a system capable of creating the aforementioned sensations. This Phase I effort involved an engineering project to design a hardware and software system capable of (a) generating tactile feedback sensations on a human operator of a Seabotix ROV hand-controller, (b) collecting real-time acceleration data from the Seabotix ROV during operation, and (c) creating real-time feedback control in which tactile sensations are automatically triggered by acceleration events detected on the ROV. In this way, the tactile feedback strategies described above could be implemented.

The overall engineering project was divided into three sub-efforts. A **tactile engine effort** in which a microcontroller, an actuator, and a set of software routines were developed to enable high fidelity tactile sensations to be imparted upon a human operator. An **ROV sensor system effort** in which an acceleration sensor was integrated with the ROV and connected over the standard ROV tether line to the microcontroller. And a **tactile control software effort** in which software routines were developed to generate tactile sensations in response to detected acceleration events upon the ROV. The status of each of these efforts are described in the following paragraphs:

a) **Tactile Engine Development**: To achieve the generation of high fidelity tactile sensations that could be modulated under software control in both magnitude and frequency, for varying time durations, a flexible system was developed using a microcontroller from Atmel Corporation (ATMega644), a high performance linear voice coil actuator from M2W Technologies (NCM02-05-005-AJB), and a motor driver from Texas Instruments (SN754410). Using electronics and code developed for this project, the tactile sensations could be imparted by the voice coil actuator at a define magnitude and frequency, for a defined duration. To match the human perceptual range, the system was configured to generate sensations between **0 to 180** HZ, with a force magnitude that could be varied from **0 to 2.25** Newtons.

b) **ROV Sensor System Development**: To detect acceleration events imparted upon the underwater vehicle from sea swells and/or collisions a three-axis accelerometer board with serial communication capabilities was purchased from Sparkfun Electronics

and integrated with the ROV. This involved mounting the accelerometer within the ROV and integrating with the Seabotix Data Board within the ROV electronics. Using this data board, acceleration data in serial packets can be transferred to the surface along the standard 500 ft teather of the ROV vehicle. This was successfully performed in a bench test. Because the ROV needs to be sealed and waterproof, we funded the manufacturer of the ROV (Seabotix) to perform final integration of the accelerometer with the data communication board. Unfortunately it was determined that the data communication board was not functioning as specified by the ROV manufacturer and was losing data packets. We plan to deliver the ROV back to the manufacturer this summer so they can fix the current problem. This unexpected problem with third party equipment is currently holding up final testing of the system.

c) **Tactile Control Software Development:** To generate tactile sensations in response to detected acceleration events, software routines were developed and integrated into the Atmel microcontroller. The software routines are configured to detect two different threshold levels of acceleration impulse upon the ROV, a first threshold level that is set to detect a sea swell event and a second higher threshold level that is set to detect a collision of the ROV with a stationary object such as the ocean floor or a pier piling. This software was successfully generated and bench-tested, but could not be tested in an underwater setting because of the ROV data communication problem as described above. We hope to have the manufacturer of the ROV fix the communication problem this summer so that we can test the software in a real-world underwater setting.

The final testing effort in which the full system was to be deployed in an underwater test and feedback is collected from marine professionals was not completed due to technical difficulties encountered with the third-party ROV electronics. After extensive interactions with the manufacturer of the ROV hardware, we determined that their ROV system was simply unable to function as advertised and thus was unable to support the testing phase of this project. That said, the Tactile Engine and ROV Sensor System were each bench-tested the system out of the water (i.e. without using the 500 ft tether) and all components operated expected. Thus although we had to abandon the project prior to the final underwater testing phase, we were able to confirm that our development electronics were able to detect acceleration events and generate tactile feedback sensations in response to those events.

TECHNICAL REPORT

Section 1.0 Introduction

1.1 Background

Remotely operated vehicles (ROV) are becoming important as military tools, research instruments, and surveillance platforms, allowing the human operators to project their sensory presence and manipulative skills into hostile environments while remaining safely protected in control rooms and other distant locations. Yet despite recent advances in feedback and control, it is still quite difficult for human operators to precisely guide remote vehicles. This is because human operators lack the same level of situational awareness that on-board pilots achieve. In general, constrained sensory feedback and limited control bandwidth between operator and vehicle significantly reduces the performance of remotely operated devices as compared to direct human presence.

Figure 1. The LBV150BE ROV from Seabotix

1.2 Motivation

While enhancing in visual, audio, and haptic feedback have increased operator skill and improved situational awareness in ROV environments; there remains a substantial reduction in ROV performance as compared to onboard piloting. One important but often neglected reason for performance reductions is the fact that the operator does not have direct physical contact with the vehicle itself and therefore has a reduced sense of vehicle response. To solve this problem, we will build a modified ROV control system that creates electronically generated engine vibration sensations and imparts these tactile sensations upon the remote operator. A number of control strategies will be developed, tested, and compared for effectiveness.

Section 2.0 Design Development

2.1 Specifications

The specifications were defined by the overall goal of our project. They are as follows:

- It is suggested to use C or C++ programming code and an ATmega644 microcontroller.
- Collisions and motion due to ocean currents, inclines, headwinds, or other factors that affect vehicle response is detected by an accelerometer then transmitted back to the microcontroller, where the data will be interpreted.
- The accelerometer unit needs to be water-proofed.
- The data from the accelerometer unit is transmitted over 500ft of water-proof cable.
- The microcontroller receives the acceleration data using RS232 asynchronies with baud rate of 9600, interpret it, and then output a proportional square wave that drives a voice coil at 1amp to produce vibrations on the ROV controller.
- A Polybot Board v1.1, developed by Dr. J. Seng, California Polytechnic State University Computer Science Professor, is used as the platform.
- The square wave will be variable in amplitude and frequency, 0-100% and 10-50 Hz, respectively.
- Seabotix LBV-150BE ROV system, an underwater ROV system currently owned by the Cal Poly Center for Coastal Marine Sciences, is used as the test vehicle.

2.1 Concept Design

2.1.1 Original Design

We started our initial design concept by following the design specifications as well as suggestions by Dr. J. Ridgely, California Polytechnic State University Mechanical Engineering Professor. The ROV will be fitted with one or more sensors, including:

- Accelerometer to measure engine vibration
- Microphone to measure audio
- Tachometer to measure engine speed
- Ammeter to measure motor power consumption

The manual control side of the system will be fitted with haptic actuators including a voice-coil driven vibro-tactile sensation generator. The system will be controlled by C++ and/or C programming code and a low-cost microcontroller, an ATmega644, which reads engine sensor data from the ROV and controls the haptic actuator on the manual control interface to produce informative haptic feedback sensations.

2.1.2 Design Update

While testing with the ROV it was determined that there was no useful data to be obtained from engine vibration, audio, velocity, or power consumption. We decided to mount the accelerometer inside the ROV and use the internal data board for transmission and pull power from the Seabotix motherboard.

With the sensor mounted inside, we can measure the major events, such as a swell or collision, and ignore other vibrations caused from the thrusters

Section 3.0 Design

3.1 Hardware Design

3.1.1 Mechanical

To achieve a vibration that could be varied by both amplitude and frequency we chose to use a linear voice coil. To drive the vibration actuator we use a SN754410 motor driver and quickly change the direction. The resulting signal is a square wave varying between +5 and –5V overlaid by a higher frequency PWM signal to control the amplitude.

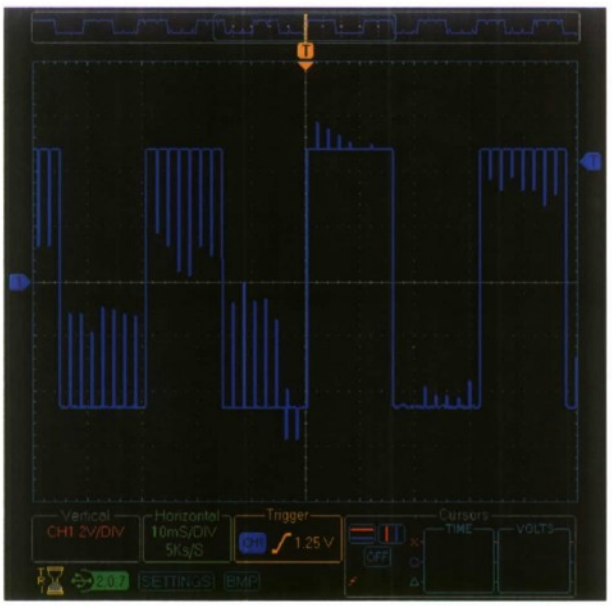


Figure 2. Signal designed to shake linear voice coil at varying amplitudes and frequencies.

To adjust the force of the vibration we attached a weight to the end of the moving shaft. The voice coil is capable of continuously moving 0.5 lbs. And has a peak force of 1.5 lbs. A mount was machined out of Delrin plastic to attach the voice coil inside the hand controller provided by Seabotix.

Figure 3. Hand controller unit provided by Seabotix.

Figure 4. Non-commutated DC linear voice coil actuator

3.1.2 Computer/Electrical

To sense the different events, such as a large swell or collision, we chose to use a 3-axis accelerometer. The sensor is mounted inside the ROV and a small Microchip processor takes that data from the accelerometer and converts it to a digital signal and sends it through the RS-232 data channel on the ROV. However, before transmitting the data we needed to correct the signal to RS-232 voltages, +10 and –10V. To accomplish this we use a MAX232 chip to shift the signal to TTL and then back to the RS-232 standards.

On the surface we pull the signal from another data board inside the power supply. This singal is then routed through another MAX232 chip to shift the RS-232 level to TTL and then fed into another processor. We chose to use an Atmega644 chip clocked at 20Mhz. The signal is brought into the microcontroller's UART and then processed.

[1] Printed from H2W Technologies, http://www.h2wtech.com/

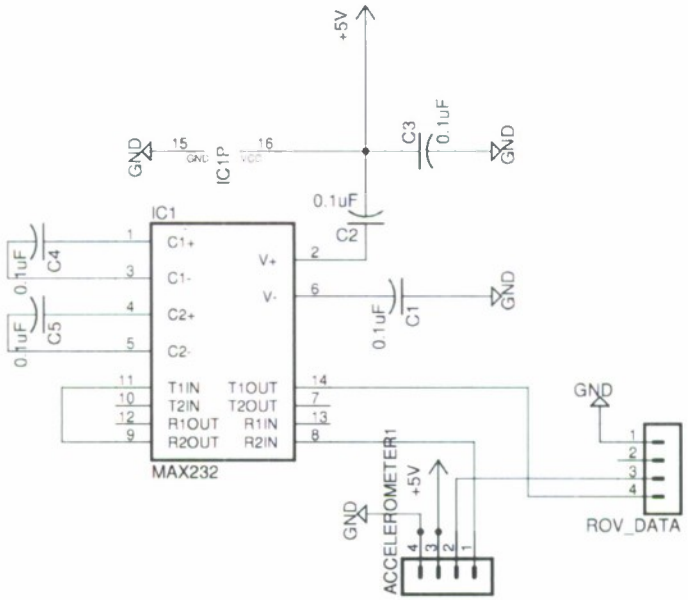


Figure 5. Schematic detailing the circuit designed to correct the accelerometer's data to real RS-232 voltages.

3.2 Software Design

3.2.1 Main Program

The main program loop switches between three tasks, one task to read the sensor data, another to react on the data, and another to send the sensor a start character in the event the accelerometer stops transmitting data.

Table 1. Summary of the tasks performed by the main program.

Main Program			
Task	**Description**	**Timing Constraint**	**States**
1	Read Accelerometer	10 micro-sec	11
2	Shake Voice Coil	10 ms	5
3	Send Start Character	50 ms	1

3.2.2 Read Accelerometer

The sensor begins each transmission of data with a '#' and '@' character. The task to read the sensor data ensures that the data is received in the correct order by looking for these characters. When they are seen the program continues reading the remaining bytes of

data. It is most important that this task is run very often, if the program is blocked from entering this code the data will be misread or the program will be stuck looking for the '#@' start condition.

Binary Output Format	
Byte	Description
0	Start Character - always #
1	Data Output Designator - always @
2	Sample Number High Byte
3	Sample Number Low Byte
4	X Axis High Byte
5	X Axis Low Byte
6	Y Axis High Byte
7	Y Axis Low Byte
8	Z Axis High Byte
9	Z Axis Low Byte
10	End Character - always $

Figure 6. Description of the data output from the accelerometer.

Once the data has been received, a '$' character is seen, the program performs a union to combine the two bytes of data for each x-, y-, and z-axes into an array. The data is then ready to be processed.

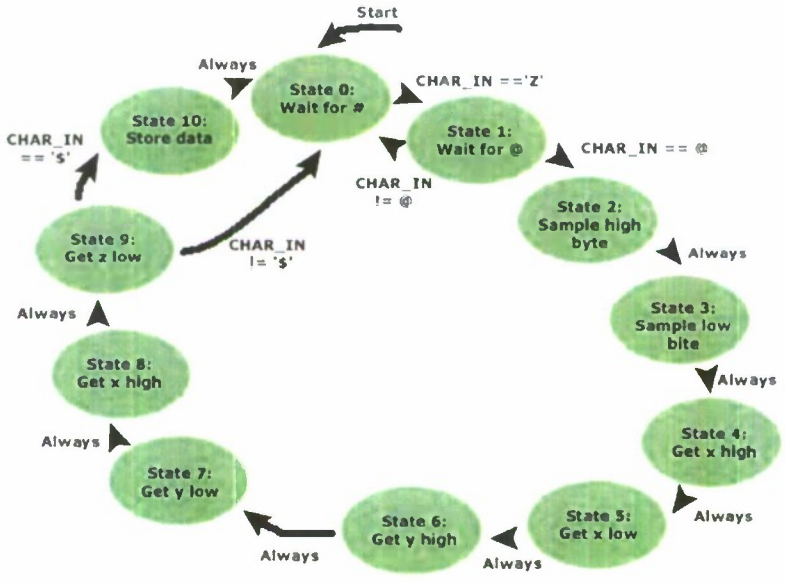
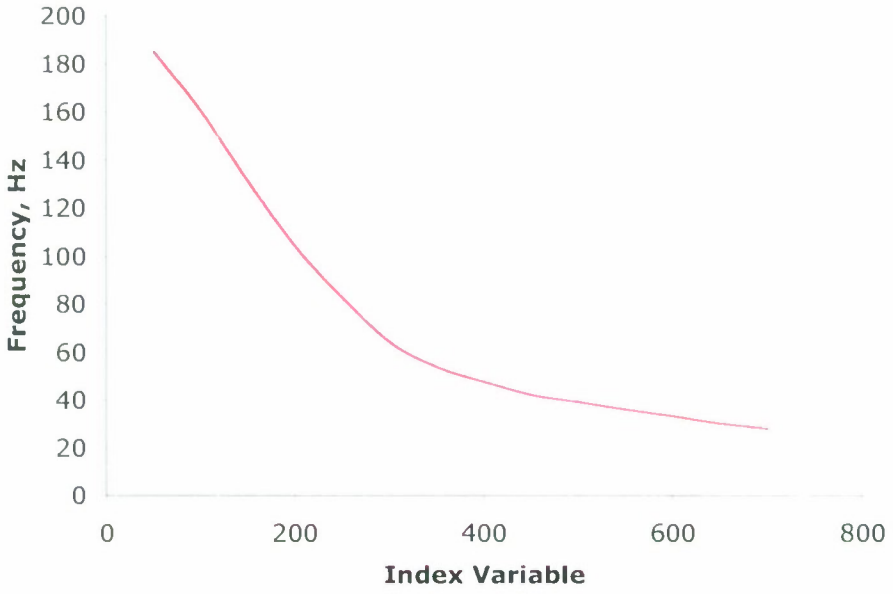


Figure 7. State transition diagram for the Read Accelerometer task.

3.2.3 Shake Voice Coil

The task to shake the voice coil first checks the arrays of data containing the acceleration information, it then makes a decision based upon these variables. The data could either indicate a swell or a collision. These two events differ in the amplitude, frequency, and duration of shaking the voice coil. When the program determines the voice coil should shake, the motor driver is set in one direction and an index variable is increased. The index variable controls the frequency, and when it reaches its maximum the motor direction is reversed. This continues until the duration limit is met.



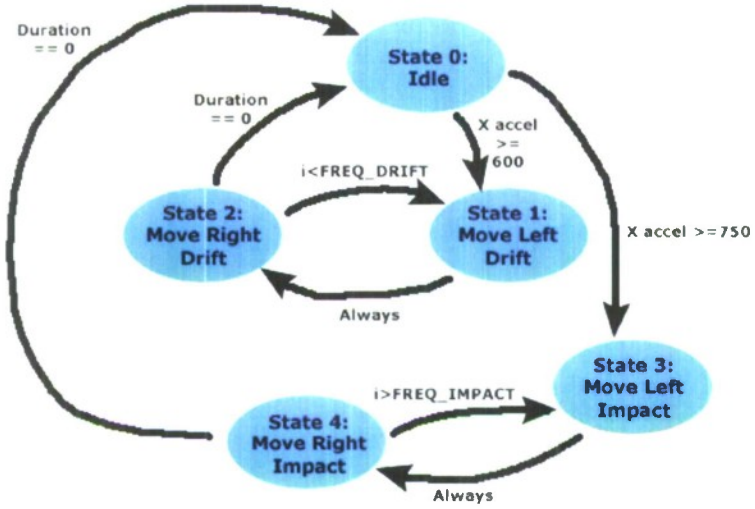


Figure 8. State transition diagram for the Voice Coil Actuation task.

3.2.4 Send Start Character

This task is entered when an index variable is reached. Its purpose is to periodically send a 'S' character to the accelerometer. This is required to start the data stream. Although it is only required at the start of the program we send it often to ensure the data does not stop streaming for any reason.

Section 3.0 Results

The prototype of the force feedback system worked perfectly. The simulated vibration sensations are produced with a magnitude and frequency that represent the real-time accelerometer behavior based on sensor readings. The only major downfall of the system was that we did not have enough time to properly tune the channel after replacing the prototype system to the real setting. Tuning the channel through the 500 ft waterproof cable was a lot more difficult than expected. We will continue working on this for the next couple of weeks.

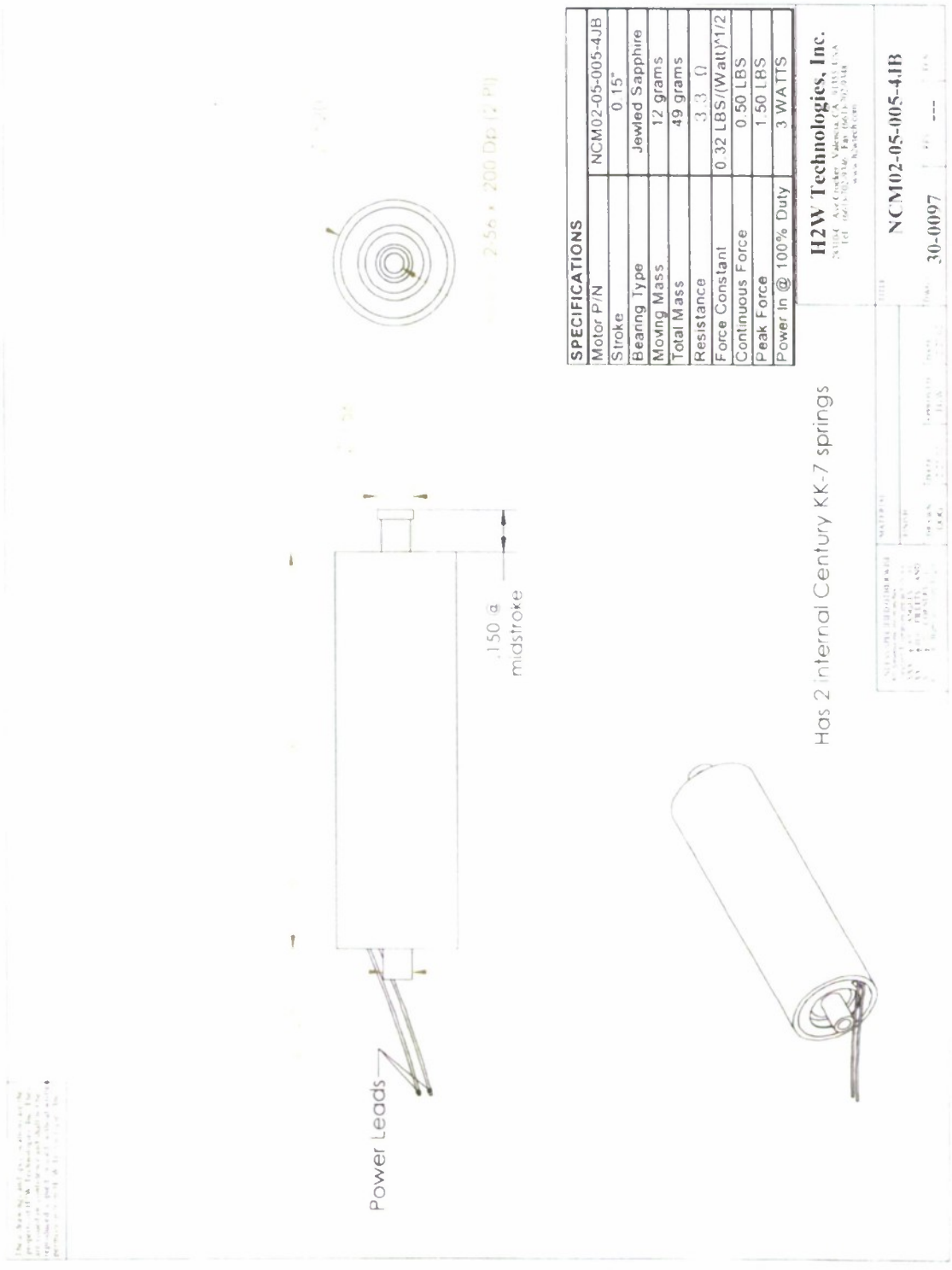
Section 4.0 Conclusion

Overall, we feel the project was a success. We noticed the voice coil produces a high frequency noise when being used. This is probably due to the fast PWM signal. Another way to provide the force feedback would be a rotary DC motor with a weighted swing arm that can oscillate to produce the vibration. We will be performing the initial functionality tests in the water. Once the initial tests are confirmed that the system is working as expected, the marine specialists will perform a simple test in an under water environment off the Cal Poly pier. They will perform a simple standardized ocean floor survey task using the ROV, with and without the sensory enhanced features developed, to find out if the force feedback system improves the overall performance of piloting an underwater vehicle.

Section 5.0 Appendices

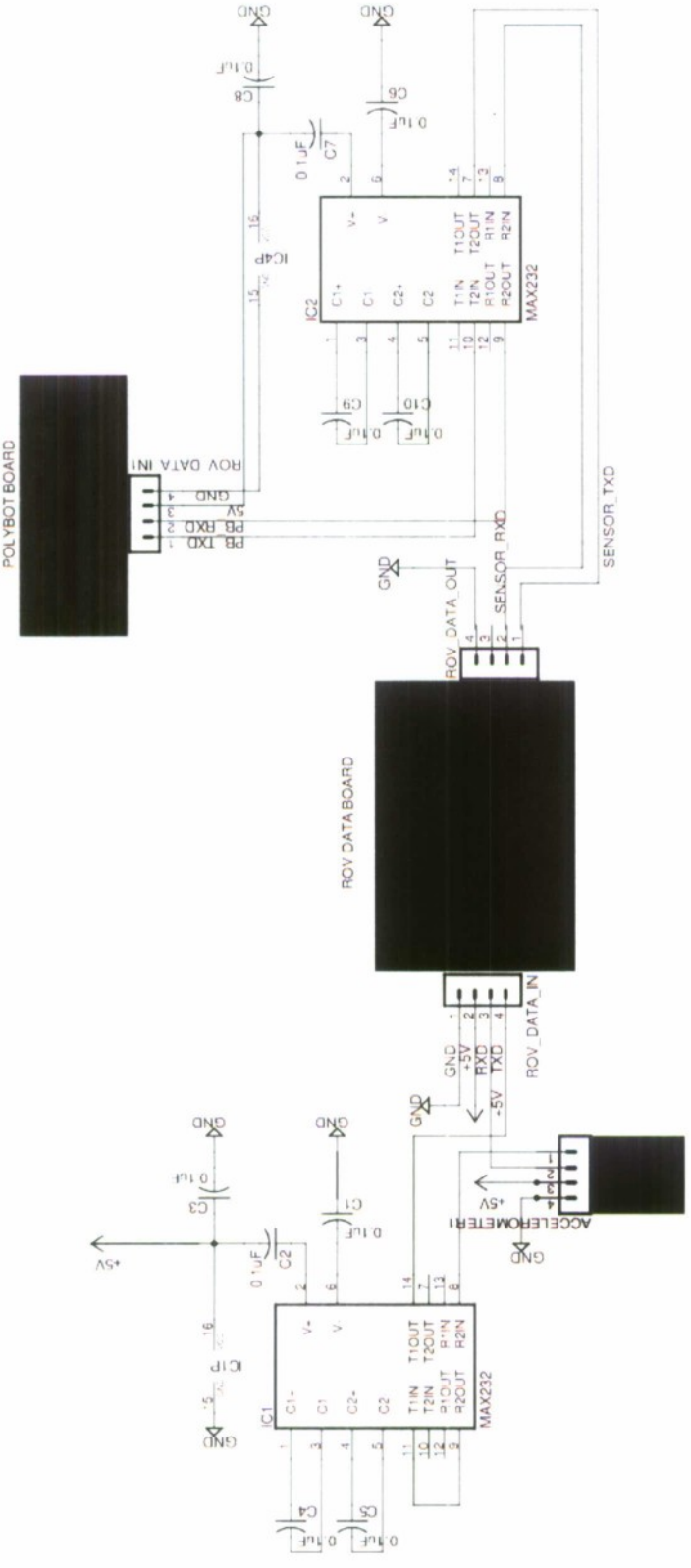
Appendix A Mechanical Drawings

Linear Voice Coil Actuator

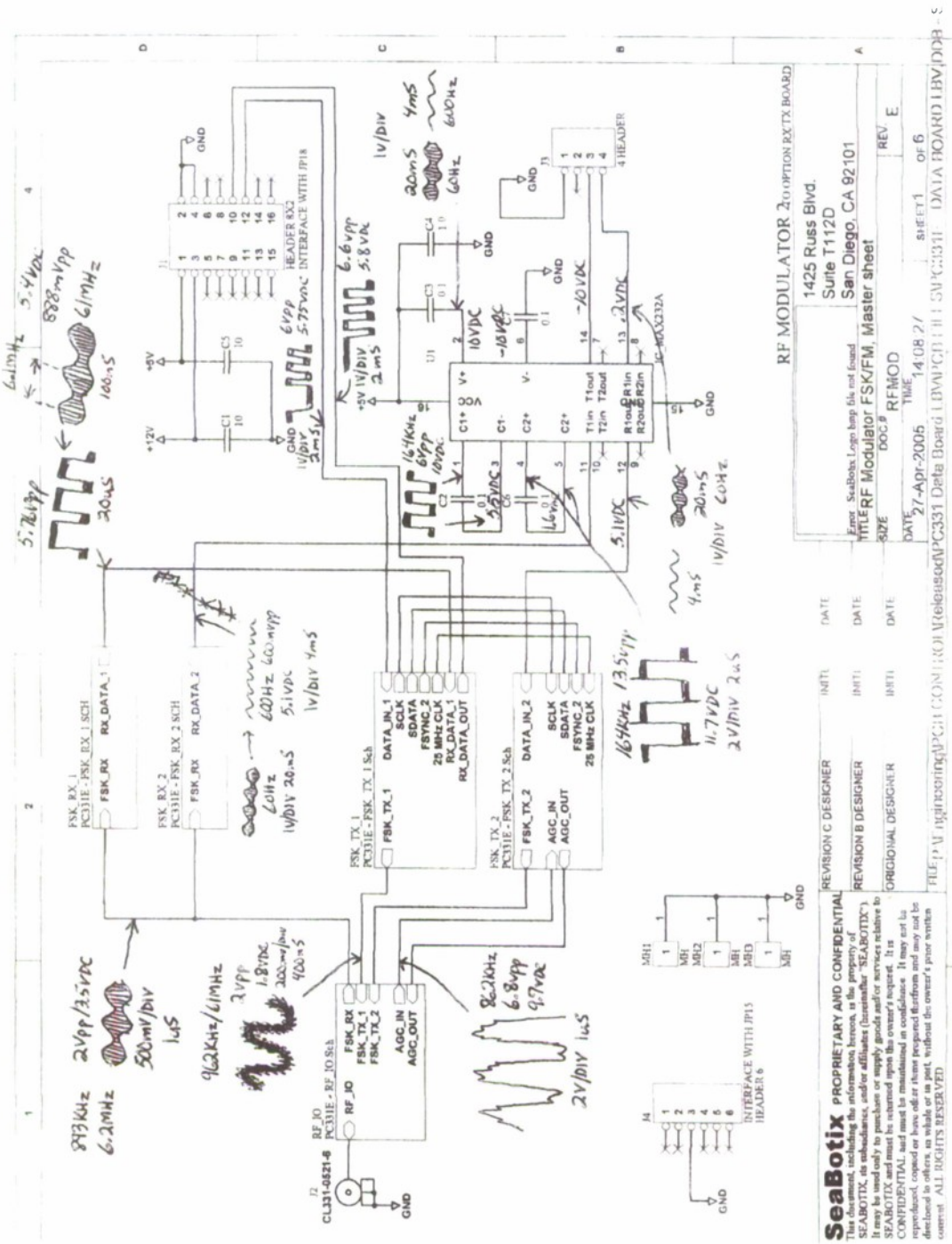


Appendix B Electronic Schematics

Electronic Design



Seabotix Data Board



Appendix C Software User's Manual

ROV_Main Data Structure Documentation

accelerometer Class Reference

`#include <Task_read_accel.h>`

Inheritance diagram for accelerometer:

IMAGE

Public Member Functions

- **accelerometer** (unsigned char)
- void **read_accel** (void)

Data Fields

- **accelerometer::u_x_axis xaxis**
- **accelerometer::u_y_axis yaxis**
- **accelerometer::u_z_axis zaxis**
- unsigned char **x_high**
- unsigned char **x_low**
- unsigned char **y_high**
- unsigned char **y_low**
- unsigned char **z_high**
- unsigned char **z_low**
- char **current_state**

Data Structures

- union **u_x_axis**
- union **u_y_axis**
- union **u_z_axis**

Detailed Description

This class contains the methods need to control the **accelerometer** read task it is a descendent of the **uart** class to allow it to use serial data transmission.

Definition at line 10 of file Task_read_accel.h.

Constructor & Destructor Documentation

accelerometer::accelerometer (unsigned char *divisor*)

The constructor method to initalize the **accelerometer** read task

Definition at line 19 of file Task_read_accel.cc.

Member Function Documentation

void accelerometer::read_accel (void)

This method controls the read **accelerometer** task. The task consists of states 0-11

State 0: check for start character of data "#"

State 1: check for Data output designator "@"

State2: receive sample number high byte, which we dont care about

State 3: reveive sample number low byte, which we dont care about

State 4: record X axis high byte

State 5: record X axis low byte

State 6: record y axis high byte

State 7: record y axis low byte

State 8: record z axis high byte

State 9: reeord z axis low byte

State 10: check for end character ($) If the end character is seen it permanently stores the data

Definition at line 27 of file Task_read_accel.cc.

References uart::check_for_char(), current_state, uart::getchar(), accelerometer::u_x_axis::x_axis_parts, x_high, x_low, accelerometer::u_y_axis::y_axis_parts, y_high, y_low, accelerometer::u_z_axis::z_axis_parts, z_high, and z_low.

Referenced by main().

The documentation for this class was generated from the following files:

- **Task_read_accel.h**
- Task_read_accel.c

accelerometer::u_x_axis Union Reference

Data Fields

- unsigned int **x_axis**
- char **x_axis_parts** [2]

Detailed Description

Definition at line 25 of file Task_read_accel.h.

The documentation for this union was generated from the following file:

- **Task_read_accel.h**

accelerometer::u_y_axis Union Reference

Data Fields

- unsigned int **y_axis**
- char **y_axis_parts** [2]

Detailed Description

Definition at line 32 of file Task_read_accel.h.

The documentation for this union was generated from the following file:

- **Task_read_accel.h**

accelerometer::u_z_axis Union Reference

Data Fields

- unsigned int **z_axis**
- char **z_axis_parts** [2]

Detailed Description

Definition at line 39 of file Task_read_accel.h.

The documentation for this union was generated from the following file:

- **Task_read_accel.h**

shake Class Reference

`#include <shake.h>`

Inheritance diagram for shake:

IMAGE

Public Member Functions

- **shake** (unsigned int)
- void **vibrate** (int, int)

Data Fields

- char **current_state**
- unsigned long **start_time**
- char **duration**
- int **impact_freq**
- int **drift_freq**
- int **impact_amp**
- int **drift_amp**
- int **i**
- int **g**
- int **on**

Detailed Description

This class contains the fuctions that control the **shake** state of the ROV Haptic system It is a decendent of the TCNT class to allow for the motor to be run on a timer.

Definition at line 10 of file shake.h.

Constructor & Destructor Documentation

shake::shake (unsigned int *prescaler*)

The constructor method to initalize the **accelerometer** read task

Definition at line 26 of file shake.cc.

References current_state, g, i, and on.

Member Function Documentation

void shake::vibrate (int *x_limit*, int *y_limit*)

This is the states that controls the shaking of the voice coil

State 0: Idle condition when the motor has no power. If the x value is greater than 650 it is experiencing a swell. If it is greater than 750 the ROV is experiencing an impact

State 1: Causes the voice coil to move left due to a drift, then transitions to state 2

State 2: Causes the voice coil to move right due to a drift, then transitions back to state 1 unless the duration is equal to zero in which case it will transition back to state 0

State 3: Causes the voice coil to move left due to a drift, then transitions to state 4

State 4: Causes the voice coil to move right due to a drift, then transitions back to state 3 unless the duration is equal to zero in which case it will transition back to state 0

Definition at line 39 of file shake.cc.

References current_state, drift_amp, drift_freq, duration, g, i, impact_amp, impact_freq, and on.

Referenced by main().

The documentation for this class was generated from the following files:

- **shake.h**
- **shake.cc**

tcnt Class Reference

`#include <Timer.h>`

Inheritance diagram for tcnt:

IMAGE

Public Member Functions

- **tcnt** (unsigned int presealer)
- void **clear_timer** (void)
- unsigned long **get_current_time** (void)
- unsigned long **start_stopwatch** (void)
- unsigned long **stop_stopwatch** (unsigned long, char)
- void **calibration** (unsigned char, unsigned char, unsigned char)
- char **hour_of_day** (void)
- char **min_of_day** (void)
- char **sec_of_day** (void)

Data Fields

- unsigned long **sw_start**
- unsigned long **time_passed**

Protected Attributes

- unsigned long **cycles_past**
- **tcnt::time_o_day day**

Data Structures

- union **time_o_day**

Detailed Description

This class controls the Timer/counter 1 on an AVR processor. The class has stop watch capabilities and can keep track to the current time of day and number of days passed.

Definition at line 53 of file Timer.h.

Constructor & Destructor Documentation

tcnt::tcnt (unsigned int *prescaler*)

this is the constructor function initializes the timer counter

Definition at line 30 of file Timer.cc.

Member Function Documentation

void tcnt::clear_timer (void)

Thismethods clears the clears the timer

Definition at line 40 of file Timer.cc.unsigned long tcnt::get_current_time (void)

This method returns the clock cylces in the 32 bit timer

Definition at line 47 of file Timer.cc.

References time_u::big_time, and time_u::time_parts.unsigned long tcnt::start_stopwatch (void)

This function starts the stop watch timer

Definition at line 55 of file Timer.cc.

References time_u::big_time, and time_u::time_parts.unsigned long tcnt::stop_stopwatch (unsigned long *start_time*, char *time_inc*)

This method returns the time difference from when start_stopwatch was started and this function was called. It can return time in either seconds, milliseconds, or micro seconds by inputting "sec", "milli_sec" or "micro_sec"

Definition at line 64 of file Timer.cc.

References time_u::big_time, and time_u::time_parts.void tcnt::calibration (unsigned char *hour*, unsigned char *min*, unsigned char *seconds*)

This method takes calabrates the Time of day to the current time when the user inputs the hour in 24 hour time, the minute, and the second

Definition at line 102 of file Timer.cc.

References tcnt::time_o_day::big_day, cycles_past, and tcnt::time_o_day::day_parts.char tcnt::hour_of_day (void)

This method returns current hours of the day. It is not how much time has passed in seconds but rather the hours portion of the total day time.

Definition at line 114 of file Timer.cc.

References time_u::big_time, and time_u::time_parts.char tcnt::min_of_day (void)

This method returns current minutes of the day. It is not how much time has passed in seconds but rather the minutes portion of the total day time.

Definition at line 129 of file Timer.cc.

References time_u::big_time, and time_u::time_parts.char tcnt::sec_of_day (void)

This method returns the current seconds of the day. It is not how much time has passed in seconds but rather the seconds portion of the total day time.

Definition at line 139 of file Timer.cc.

References time_u::big_time, and time_u::time_parts.

The documentation for this class was generated from the following files:

- Timer.h
- Timer.cc

tcnt::time_o_day Union Reference

Data Fields

- unsigned long **big_day**
- int **day_parts** [2]

Detailed Description

Definition at line 66 of file Timer.h.

The documentation for this union was generated from the following file:

- Timer.h

time_u Union Reference

Data Fields

- long **big_time**
- int **time_parts** [2]

Detailed Description

Definition at line 44 of file Timer.h.

The documentation for this union was generated from the following file:

- Timer.h

uart Class Reference

`#include <avr_serial.h>`

Inheritance diagram for uart:

IMAGE

Public Member Functions

- **uart** (unsigned char)
- char **putchar** (char)
- void **puts** (char *)
- char **check_for_char** (void)
- char **getchar** (void)
- void **write_bin** (unsigned char)
- void **write_hex** (unsigned char)
- void **write** (unsigned char)
- void **write** (char num)
- void **write_bin** (unsigned int)
- void **write_hex** (unsigned int)
- void **write** (unsigned int)
- void **write** (int)
- void **write_hex** (unsigned long)
- void **write** (long)

Protected Member Functions

- char **transmitter_empty** (void)

Detailed Description

This class controls a UART (Universal Asynchronous Receiver Transmitter), a common serial interface. It talks to old-style RS232 serial ports (through a voltage converter chip such as a MAX232) or through a USB to serial converter such as a FT232RL chip. The UART is also sometimes used to communicate directly with other microcontrollers, sensors, or wireless modems.

This class has originally been written for AVR processors which have only one UART, but it should be extendable for use with processors which have dual UARTs.

Definition at line 206 of file avr_serial.h.

Constructor & Destructor Documentation

uart::uart (unsigned char *divisor*)

This method sets up the AVR UART for communications. It enables the appropriate inputs and outputs and sets the baud rate divisor.

Parameters:

divisor The baud rate divisor to be used for controlling the rate of communication. See the *.h file in which various values of the divisor are defined as macros.

Definition at line 37 of file avr_serial.cc.

Member Function Documentation

char uart::transmitter_empty (void) [protected]

This function checks if the serial port transmitter is ready to send data. It simply tests the bit in the serial port status register which indicates that the transmitter buffer is empty.

Returns:

0 if the transmitter is empty and ready to send, and 1 if not

Definition at line 87 of file avr_serial.cc.char uart::putchar (char *chout*)

This method sends one character to the serial port. It waits until the port is ready, so it can hold up the system for a while. It times out if it waits too long to send the character; you can check the return value to see if the charaeter was sueeessfully sent, or just cross your fingers and ignore the return value. Note: It's possible that at slower baud rates and/or higher proeessor speeds, this routine might time out even when the port is working fine. A solution would be to change the count variable to an integer and use a larger starting number. Note 2: Fixed! The count is now an integer and it works at lower baud rates.

Parameters:

chout The character to be sent out

Returns:

0 if everything was OK and (char)(-1) if there was a timeout

Definition at line 67 of file avr_serial.cc.

Referenced by puts(), write_bin(), and write_hex().void uart::puts (char * *str*)

This is the usual...it just writes all the characters in a string until it gets to the '\0' at the end. Warning: By repeatedly calling **putchar()**, this method can hold up the program while it's running, and so it shouldn't be used when the program has to meet timing constraints.

Parameters:

str The string to be written

Definition at line 104 of file avr_serial.cc.

References putchar().

Referenced by main(), and write().char uart::check_for_char (void)

This funetion cheeks if there is a character in the serial port's receiver buffer. It returns 1 if there's a charaeter available, and 0 if not.

Returns:

1 for character available, 0 for no character available

Definition at line 134 of file avr_serial.cc.

Referenced by accelerometer::read_accel().char uart::getchar (void)

This method gets one character from the serial port, if one is there. If not, it waits until there is a character available. This can sometimes take a long time (even forever), so use this function carefully. It's generally safer to use **check_for_char()** to ensure that there's data available first.

Returns:

The character which was found in the serial port receive buffer

Definition at line 118 of file avr_serial.cc.

Referenced by accelerometer::read_accel().void uart::write_bin (unsigned char *num*)

This method writes a character to the serial port as a text string showing the 8-bit unsigned number in that character in binary form.

Parameters:

num The 8-bit number to be sent out

Definition at line 149 of file avr_serial.cc.

References putchar().void uart::write_hex (unsigned char *num*)

This method writes a character to the serial port as a text string showing the 8-bit unsigned number in that character in hexadecimal form.

Parameters:

num The 8-bit number to be sent out

Definition at line 167 of file avr_serial.cc.

References putchar().

Referenced by write_hex().void uart::write (unsigned char *num*)

This method writes a character to the serial port as a text string showing the 8-bit unsigned number in that character.

Parameters:

num The 8-bit number to be sent out

Definition at line 191 of file avr_serial.cc.

References puts().void uart::write (char *num*)

This method writes a character to the serial port as a text string showing the 8-bit signed number in that character.

Parameters:

num The 8-bit number to be sent out

Definition at line 205 of file avr_serial.cc.

References puts().void uart::write_bin (unsigned int *num*)

This method writes an integer to the serial port as a text string showing the 16-bit unsigned number in that character in binary form.

Parameters:

num The 16-bit number to be sent out

Definition at line 219 of file avr_serial.cc.

References putchar().void uart::write_hex (unsigned int *num*)

This method writes an integer to the serial port as a text string showing the 16-bit unsigned number in that integer in hexadecimal notation.

Parameters:

num The 16-bit number to be sent out

Definition at line 237 of file avr_serial.cc.

References putchar().void uart::write (unsigned int *num*)

This method writes an integer to the serial port as a text string showing the 16-bit unsigned number in that integer.

Parameters:

num The 16-bit number to be sent out

Definition at line 261 of file avr_serial.cc.

References puts().void uart::write (int *num*)

This method writes an integer to the serial port as a text string showing the 16-bit signed number in that integer.

Parameters:

num The 16-bit number to be sent out

Definition at line 275 of file avr_serial.cc.

References puts().void uart::write_hex (unsigned long *num*)

This method writes a long integer to the serial port as a text string showing the 32-bit unsigned number in that long integer.

Parameters:

num The 32-bit number to be sent out

Definition at line 304 of file avr_serial.cc.

References write_hex().void uart::write (long *num*)

This method writes a long integer to the serial port as a text string showing the 32-bit signed number in that long integer.

Parameters:

num The 32-bit number to be sent out

Definition at line 289 of file avr_serial.cc.

References puts().

The documentation for this class was generated from the following files:

- **avr_serial.h**
- **avr_serial.cc**

ROV_Main File Documentation

avr_serial.cc File Reference

```
#include <stdlib.h>
#include <avr/io.h>
#include <avr/interrupt.h>
#include "avr_serial.h"
```

Detailed Description

This file contains functions which allow the use of a serial port on an AVR microcontroller. Compatibility macros are provided to isolate the names of various registers from the many specific AVR device types.

This code is designed to work for low-performance applications without requiring the use of interrupts. Interrupt based receiving code has not been completed or tested.

Revised:

- 04-03-06 JRR For updated version of compiler
- 06-10-06 JRR Ported from C++ to C for use with some C-only projects; also the serial_avr.h header has been stuffed with defines for compatibility among lots of AVR variants
- 08-11-06 JRR Some bug fixes
- 03-02-07 JRR Ported back to C++. I've had it with the limitations of C.

Definition in file **avr_serial.cc**.

avr_serial.h File Reference

Data Structures

- class **uart**

Detailed Description

This file contains functions which allow the use of a serial port on an AVR microcontroller. Compatibility macros are provided to isolate the names of various registers from the many specific AVR device types.

This code is designed to work for low-performance applications without requiring the use of interrupts. Interrupt based receiving code has not been completed or tested.

Revised:

- 04-03-06 JRR For updated version of compiler
- 06-10-06 JRR Ported from C++ to C for use with some C-only projects; also the serial_avr.h header has been stuffed with defines for compatibility among lots of AVR variants
- 08-11-06 JRR Some bug fixes
- 03-02-07 JRR Ported back to C++. I've had it with the limitations of C.

Definition in file **avr_serial.h**.

rov_main.cc File Reference

```
#include <stdlib.h>
#include <avr/io.h>
#include <avr/interrupt.h>
#include <math.h>
#include <util/delay.h>
#include "globals.h"
#include "Timer.h"
#include "avr_serial.h"
#include "Task_read_accel.h"
#include "polybot.h"
#include "shake.h"
```

Defines

- #define **MY_DIVISOR** 129

Functions

- int **main** (void)

Detailed Description

This program receives serial data in from a sparkfun serial **accelerometer** v.5 in Binary mode written by: WCK and DG Revisions:

Definition in file **rov_main.cc**.

Define Documentation

#define MY_DIVISOR 129

This is the baud rate divisor for the UART. Values which have worked: 25: 9600 baud, 4 MHz erystal ose., works with FT232RL RS232-USB chip 51: 4800 baud 103: 2400 baud

Definition at line 31 of file rov_main.cc.

Referenced by main().

Function Documentation

int main (void)

This main function creates a serial port object, then receives the binary serial data from the serial **accelerometer**.

Returns:

The return value from **main()** doesn't mean anything in a microcontroller, because the main function should never be exited. There is no operating system to which to return. However, we put a return value here anyway because the C/C++ language standard expects it, and the compiler will spit out warnings if it's not there.

Definition at line 44 of file rov_main.cc.

References accelerometer::current_state, shake::drift_amp, shake::drift_freq, shake::impact_amp, shake::impact_freq, MY_DIVISOR, uart::puts(), accelerometer::read_accel(), shake::vibrate(), accelerometer::u_x_axis::x_axis, accelerometer::xaxis, accelerometer::u_y_axis::y_axis, accelerometer::yaxis, accelerometer::u_z_axis::z_axis, and accelerometer::zaxis.

shake.cc File Reference

```
#include <stdlib.h>
#include <avr/io.h>
#include <avr/interrupt.h>
#include <util/delay.h>
#include "globals.h"
#include "Timer.h"
#include "shake.h"
#include "polybot.h"
```

Detailed Description

This file contains the funtions for the **shake** task Written by: WCK and DG

Revised:

Definition in file **shake.cc**.

shake.h File Reference

Data Structures

- class **shake**

Detailed Description

This class controls the **shake** task for a voice coil. Written By: WCK and DG

Definition in file **shake.h**.

Task_read_accel.h File Reference

Data Structures

- class **accelerometer**
- union **accelerometer::u_x_axis**
- union **accelerometer::u_y_axis**
- union **accelerometer::u_z_axis**

Detailed Description

This class controls the read **accelerometer** task. The **accelerometer** is a Sparkfun SerAccel V5

Definition in file **Task_read_accel.**

Appendix D Source Code

Main Control Program

```
//--------------------------------------------------------------------------------

/** \file  rov_main.cc
 *
 *  This program receives serial data in from a sparkfun serial accelerometer v.5
 *  in Binary mode.
 *
 *   written by: WCK and DG
 *  Revisions:
 *
 */

#include <stdlib.h>
#include <avr/io.h>
#include <avr/interrupt.h>
#include <math.h>

extern "C" {
#include <util/delay.h>
#include "globals.h
}

#include "Timer.h"
#include "avr_serial.h"
#include "Task_read_accel.h"
#include "polybot.h"
#include "shake.h"

/** This is the baud rate divisor for the UART.  Values which have worked:
 *   25: 9600 baud, 4 MHz crystal osc., works with FT232RL RS232-USB chip
 *   51: 4800 baud
 *   103: 2400 baud
 */

#define MY_DIVISOR         129

//--------------------------------------------------------------------------

/** This main function creates a serial port object, then receives the binary serial
*      data from the serial accelerometer.
*
*  @return The return value from main() doesn't mean anything in a microcontroller,
*          because the main function should never be exited. There is no operating
*          system to which to return. However, we put a return value here anyway
*          because the C/C++ language standard expects it, and the compiler will
*          spit out warnings if it's not there.
*/

 int main (void)
        {
        unsigned char char_in;              // Character typed by user
        char reading;
        char entry;
        char send_s=0;

        //Calls the Constructors for Accelerometer read task and the Shake task
         accelerometer ser_accel (MY_DIVISOR);
        shake vc_shake (0);

        // Initialization functions for the Polybot board drivers
        initialize();
        motor_init();

        // defines the amplitude and frequency for the impact and drift
        vc_shake.impact_freq = 400;
        vc_shake.drift_freq = 800;
        vc_shake.impact_amp = 90;
        vc_shake.drift_amp = 45;
```

```
// Clears the values in the X, Y, and Z memory locations
ser_accel.xaxis.x_axis=0;
ser_accel.yaxis.y_axis=0;
ser_accel.zaxis.z_axis=0;
ser_accel.current_state = 0;
ser_accel.puts ("S");                //  Requests data from the Accelerometer

// main while loop of the program
while (1)
{
        ser_accel.read_accel();
        unsigned int temp_x = ser_accel.xaxis.x_axis / 100;
        unsigned int temp_y = ser_accel.yaxis.y_axis / 100;
        vc_shake.vibrate (temp_x, temp_y);    // runs the shake task

        // Sends out a consistent signal requesting data from the accelerometer
        send_s++;

        if(send_s >= 50)
           {
           send_s=0;
           ser_accel.puts("S");
           }
        }
}
```

Read Accelerometer Task Header File

```
//---------------------------------------------------------------------------

/** \file Task_read_accel.h This class controls the
read accelerometer task.  The accelerometer
*       is a Sparkfun SerAccel  V5
*
*/


/** This class contains the methods need to control the accelerometer read task
*       it is a descendent of the uart class to allow it to use serial data transmission.
*/

class accelerometer : public uart
    {
    // Private data and methods are accessible only from within this class and
    // cannot be accessed from outside -- even from descendents of this class
    private:

    // Protected data and methods are accessible from this class and its descendents
    // only
    protected:

    // Public methods can be called from anywhere in the program where there is a
    // pointer or reference to an object of this class
    public:

    // a union used to combine the low and high bytes for the x axis coming
       in from the accelerometer

            union u_x_axis
                    {
                    unsigned int x_axis;
                    char x_axis_parts[2];
                    } xaxis;


        // a union used to combine the low and high bytes for the y axis coming in
            from the accelerometer

            union u_y_axis
                    {
                    unsigned int y_axis;
                    char y_axis_parts[2];
                    } yaxis;



        // a union used to combine the low and high bytes for thez axis coming in from
          the accelerometer

            union u_z_axis
                    {
                    unsigned int z_axis;
                    char z_axis_parts[2];
                    } zaxis;

            unsigned char x_high;
            unsigned char x_low;
            unsigned char y_high;
            unsigned char y_low;
            unsigned char z_high;
            unsigned char z_low;
            char current_state;
            accelerometer (unsigned char);
            void read_accel (void);

        };
```

Read Accelerometer Task

```
//-------------------------------------------------------------------------------------

//*************************************************************************************
/** \file avr_serial.cc
*       This file contains the funtions for the accelerometer read task
*       Written by: WCK
*
*  Revised:
*/
//*************************************************************************************

#include <stdlib.h>
#include <avr/io.h>
#include <avr/interrupt.h>
#include "avr_serial.h"
#include "Task_read_accel.h"

//-------------------------------------------------------------------------------------
/** The constructor method to initalize the accelerometer read task
*/

accelerometer::accelerometer (unsigned char divisor)
        :uart (divisor)
        {
        }

//-------------------------------------------------------------------------------------
/** This method controls the read accelerometer task.  The task consists of states 0-11
*/

void accelerometer::read_accel (void)
        {
        char entry = check_for_char ();

        if (entry)
                {
                char char_in = getchar ();  //Read the character

                switch (current_state)
                        {

//---------------------------- State 0: Waiting for # -------------------------------
/** State 0: check for start character of data "#"
*/

                        case 0:
                                if (char_in == '#')
                                        {
                                        current_state ++;
                                        }
                                break;



//---------------------------- State 1: wait for @ ----------------------------------
/**State 1: check for Data output designator "@"
*/

                        case 1:
                                if (char_in == '@')
                                        {
                                        current_state ++;
                                        }
                                else
                                        {
                                        current_state = 0;
                                        }
                                break;
```

```
//--------------------------- State 2: get sample number high byte ---------------------
/** State2: receive sample number high byte, which we dont care about
*/

                        case 2:
                                current_state ++;
                                break;


//--------------------------- State 3: get sample number low byte  ---------------------
/** State 3: reveive sample number low byte, which we dont care about
*/

                        case 3:
                                current_state ++;
                                break;


//--------------------------- State 4: get x axis high byte  ---------------------------
/** State 4: record X axis high byte
*/

                        case 4:
                                x_high = char_in;
                                current_state ++;
                                break;

//--------------------------- State 5: get x axis low byte  ----------------------------
/** State 5: record X axis low byte
*/

                        case 5:
                                x_low = char_in;
                                current_state ++;
                                break;

//--------------------------- State 6: get y axis high byte  ---------------------------
/** State 6: record y axis high byte
*/

                        case 6:
                                y_high = char_in;
                                current_state ++;
                                break;

//--------------------------- State 7: get y axis low byte  ----------------------------
/** State 7: record y axis low byte
*/

                        case 7:
                                y_low = char_in;
                                current_state ++;
                                break;

//--------------------------- State 8: get z axis high byte  ---------------------------
/** State 8: record z axis high byte
*/

                        case 8:
                                z_high = char_in;
                                current_state ++;
                                break;

//--------------------------- State 9: get z axis low byte  ----------------------------
/** State 9: record z axis low byte
*/

                        case 9:
                                z_low = char_in;
                                current_state ++;
                                break;

//--------------------------- State 10: check  for end character  ----------------------
```

```
/**State 10: check for end character ($) If the end character is seen it
 *permanently stores the data
 */

                        case 10:
                                if (char_in == '$')
                                        {
                                        xaxis.x_axis_parts[1] = 0;
                                        xaxis.x_axis_parts[2] = 0;
                                        yaxis.y_axis_parts[1] = 0;
                                        yaxis.y_axis_parts[2] = 0;
                                        zaxis.z_axis_parts[1] = 0;
                                        zaxis.z_axis_parts[2] = 0;
                                        xaxis.x_axis_parts[1] = x_low;
                                        xaxis.x_axis_parts[2] = x_high;
                                        yaxis.y_axis_parts[1] = y_low;
                                        yaxis.y_axis_parts[2] = y_high;
                                        zaxis.z_axis_parts[1] = z_low;
                                        zaxis.z_axis_parts[2] = z_high;
                                        }
                                else
                                        {
                                        current_state = 0;
                                        }
                                break;
                        }
                }
        }
```

Voice Coil Actuation Task Header File

```
//-------------------------------------------------------------------------------------

/** \file shake.h
*       This class controls the shake task for a voice coil.
*       Written By: WCK and DG
*/

 /**This class contains the fuctions that control the shake state of the ROV
   *Haptic system
   *It is a decendent of the TCNT class to allow for the motor to be run on a timer.
   */

class shake : public tcnt
        {
    // Private data and methods are accessible only from within this class and
    // cannot be accessed from outside -- even from descendents of this class
    private:

    // Protected data and methods are accessible from this class and its descendents
    // only
    protected:

    // Public methods can be called from anywhere in the program where there is a
    // pointer or reference to an object of this class
    public:
                char current_state;
                unsigned long start_time;
                char duration;
                int impact_freq;
                int drift_freq;
                int impact_amp;
                int drift_amp;
                int i;
                int g;
                int on;
                shake (unsigned int);
                void vibrate (int, int);
        };
```

Voice Coil Actuation Task

```
//-------------------------------------------------------------------------------------

//*************************************************************************************
/** \file shake.cc
*       This file contains the funtions for the shake task
*       Written by: WCK and DG
*
*  Revised:
*/
//*************************************************************************************

#include <stdlib.h>
#include <avr/io.h>
#include <avr/interrupt.h>

extern "C" {
#include <util/delay.h>
#include "globals.h"
}

#include "Timer.h"
#include "shake.h"
#include "polybot.h"


//-------------------------------------------------------------------------------------
/** The constructor method to initalize the accelerometer read task
*/

shake::shake (unsigned int prescaler)
        :tcnt (prescaler)
        {
        current_state = 0;
        i=0;
        g=0;
        on = 0;
        }

//-------------------------------------------------------------------------------------
/** This is the states that controls the shaking of the voice coil
*/

void shake::vibrate (int x_limit, int y_limit)
        {
        switch (current_state)
                {

//------------------------ State 0: Idle-----------------------------------------
/**  State 0: Idle condition when the motor has no power.  If the x value is greater
  *  than 650 it is experiencing a swell.
  *     If it is greater than 750 the ROV is experiencing an impact
  */
                case 0:
                        if (x_limit >= 650 || y_limit >= 650)
                                {
                                if (x_limit >= 750 || y_limit >= 750)
                                        {
                                        current_state = 3;
                                        duration = 200;
                                        }
                                else
                                        {
                                        current_state ++;
                                        duration = 200;
                                        }
                                }
                        else
                                {
                                set_motor_power(0, 0);
                                }
                        break;
```

```
//------------------------ State 1: Shake Left-----------------------------------------
/** State 1: Causes the voice coil to move left due to a drift, then transitions to
  *state 2
  */
              case 1:
                     if (i >= drift_freq)
                            {
                            i=0;
                            on = 0;
                            current_state ++;
                            set_motor_power(1,0);
                            }
                     if (i < drift_freq && on == 1)
                            {
                            i++;
                            }
                     else
                            {
                            set_motor_power(1,drift_amp);
                            on = 1;
                            i++;
                            }
                     break;

//------------------------ State 2: Shake Right----------------------------------------
/** State 2: Causes the voice coil to move right due to a drift, then transitions back
  *to state 1 unless the duration
  *     is equal to zero in which case it will transition back to state 0
  */
              case 2:
                     if (duration == 0)
                            {
                            current_state = 0;
                            }
                     else
                            {
                            if (g >= drift_freq)
                                   {
                                   g=0;
                                   on = 0;
                                   current_state = 1;
                                   set_motor_power (1,0);
                                   duration --;
                                   }

                            if (g < drift_freq && on == 1)
                                   {
                                   g++;
                                   }
                            else
                                   {
                                   set_motor_power(1,-drift_amp);
                                   on = 1;
                                   g++;
                                   }
                            }
                     break;

//------------------------ State 3: Shake Left Due to Impact----------------------------
/** State 3: Causes the voice coil to move left due to a drift, then transitions to
  * state 4
  */
              case 3:
                     if (i >= impact_freq)
                            {
                            i=0;
                            current_state ++;
                            }
                     if (i < impact_freq && on == 1)
                            {
                            i++;
                            }
```

```
                        else
                                {
                                set_motor_power(1,impact_amp);
                                on = 1;
                                i++;
                                }
                        break;

//------------------------ State 4: Shake Right Due to Impact-----------------------------
/** State 4: Causes the voice coil to move right due to a drift, then transitions
  * back to state 3 unless the duration
  *     is equal to zero in which case it will transition back to state 0
  */

                case 4:
                        if (duration == 0)
                                {
                                current_state = 0;
                                }
                        else
                                {
                                if (g >= impact_freq)
                                        {
                                        g=0;
                                        on = 0;
                                        current_state = 3;
                                        set_motor_power (1,0);
                                        duration --;
                                        }
                                if (g < impact_freq && on == 1)
                                        {
                                        g++;
                                        }
                                else
                                        {
                                        set_motor_power(1,-impact_amp);
                                        on = 1;
                                        g++;
                                        }
                                }
                        break;
                }
        }
```

Header File to Prototype Polybot Library

```
//-------------------------------------------------------------------------------------


extern "C"
{
//Functions from lcd.c
void write_lcd(u08 data);
void write_control(u08 data);
void write_data(u08 data);
void lcd_init(void);
void print_string(char* string);
void print_int(u16 number);
void print_fp(float number);
void clear_screen(void);
void next_line(void);
void lcd_cursor(u08 row, u08 col);

//Functions from motor.
void set_motor_power(u08 num, signed char direction);
void motor_init(void);
void TIMER1_COMPB_vect(void);

//Functions from utility.c
void initialize(void);
void led_on(void);
void led_off(void);
u08 get_sw1(void);
void relay_on(void);
void relay_off(void);
u08 digital(u08 num);

//Functions from delays.c
void delay_ms(u16 num);
void delay_us(u16 num);

}
```

Enhancing Human Physical Performance through Detailed Human Modeling

Project Investigator:

Peter Schuster
Mechanical Engineering
California Polytechnic State University
San Luis Obispo, CA

Abstract

Well-designed physical training regimens are increasingly critical as soldiers and athletes strive for greater levels of physical capability. Similarly, exercise protocols are crucial for post-surgical recovery and during long-duration bed-rest or spaceflight. Current training procedures are based on extensive experience, but cannot be adjusted for individuals since the specific link between physical activities and strength gain is not clear, particularly for bone. As a result, what makes one person stronger may injure another.

Physiologists know that bone growth occurs at locations of higher dynamic strain. The long-term goal of the Cal Poly Exercise Biomechanics Research Group is to assess local bone growth during specific exercises, based on numerical prediction of the dynamic strain. This will enable improved exercise protocols. In the current C3RP project, methodologies were developed for creating detailed finite element (FE) models from computed tomography (CT) scans and for predicting muscle forces during specific exercise. These are two critical steps on the path linking physical activity to predicted bone growth. The project fully met its stated goals, although final procedure testing is still in progress. Major tasks included: reviewing the literature, evaluating software, obtaining test samples, collecting data, and developing and testing the final procedures. As a result, the defined procedures reflect the state-of-the art in human modeling and biomechanical simulations.

Relevance

Athletes and soldiers are ealled upon to perform physical feats well beyond the skills of the average person. Indeed, these activities often far exceed the performer's original capabilities before extensive training. As a result, a person ineorrectly prepared for these tasks may sustain injuries sueh as sprained muscles, fraetured bones, or torn ligaments. These can lead not only to loss of the objeetive, but also greater risk of further injury to themselves and others. Current physical training proeedures to prepare for these activities have evolved over centuries. But, there is a need to achieve greater levels of performanee in shorter periods of time, as deseribed in "physieal performanee enhancement" in the Human Performanee, Training and Survivability thrust area in ONR's Code 30 Science and Teehnology Program.

Optimal exercise protocols are also critical for maintaining musculoskeletal health in long-duration spaee flights. Even with the use of speeified exercise regimens during recent missions, museuloskeletal strength ehanges have remained a eoncern, especially in the lower limb. A method is needed to determine which exercise routines provide the most time-efficient program for minimizing bone and musele loss in redueed gravity.

A third area of relevance for this work relates to the aging U.S. population. The surgeon general's report on bone health [U.S. DHHS 2004] states that osteoporosis currently affects more than 25 million Americans, and is espeeially prevalent in post-menopausal women. By 2020, half of Amerieans over the age of fifty will have or be at risk of having osteoporosis of the hip. The disease is the most eommon eause of bone fraeture, resulting in 1.5 million fractures per year in the U.S. Exercise is a critical component in preventing and treating osteoporosis. However, Kemmler, *et al.* (2004) found that "the isolated effeet of habitual physical activity, unspeeifie exercise partieipation, and musele strength on bone parameters is rather low." Choosing the right exercises is as important as exercise in general.

Introduction

Recent advances in numerical biomechanics have made it possible to develop detailed, subject-speeifie numerical models of the musculoskeletal system based on computed tomography (CT) or magnetic resonance imaging (MRI) scans. Combined with correet material properties, these models ean be used to predict the injury tolerance of a specific individual under a variety of loading eonditions. A next logical step is to use these models to specify an aggressive training program to inerease the subject's injury tolerance for specific desired aetivities.

Using this approaeh, the Cal Poly Exereise Biomechanics Research Group has embarked on a plan of research to enable physical trainers to identify speeifie exercise protocols to deliver desired results for an individual, speeifieally in terms of local bone strength. This plan includes four major steps:

- Develop detailed subject-specifie numerieal models for non-intrusive medical scans
- Predict internal museuloskeletal forees during specific exercises
- Prediet bone strains and resulting bone growth during these exercises
- Evaluate bone strength gains resulting from predieted bone growth.

The first two steps were the focus of the eurrent C^3RP-funded projeet.

Although it is widely accepted and has been demonstrated in long-term controlled trials that exercise fulfills a critical role in promoting musculoskeletal health [Cussler *et al.* 2005, Engelke *et al.* 2006], the relationship between a particular exercise and the specific benefits to a related muscle or bone is not established. Lorentzon, *et al.* (2005) reported, "neither the lowest effective amount of [physical activity] needed to induce an osteogenic response nor its effect on the BMD [bone mineral density] and size of the different bone compartments ... has yet been clarified." Studies show that specific exercises are needed to deliver the desired musculoskeletal effects identified in controlled trials. But, how can we determine which exercises to perform?

Several recent studies have attempted to answer part of this question. Miller (2004) evaluated the relative benefits of aerobic exercise and strength training in a review of thirteen prior studies. Kato *et al.* (2006) found that low repetition, high-impact training resulted in higher BMD gains than other forms of exercise. Vainionpää *et al.* (2006) also found a positive relationship between exercise intensity and BMD. Similar studies link muscle strain to musele mass gains. Stimulation of specific muscle and bone regions through mechanical loading serves as the body's steering mechanism for where and how new minerals and muscle cells are added.

Although each of these studies identifies a direction of improvement, they do not isolate specific exercises. Indeed, studying this problem with human trials is problematic since subjects would need to have very strict controls on types of exercise performed, and such controls may have a detrimental affect on health. However, combining short-duration human subject exercise testing with a validated numerical simulation procedure that can predict musculoskeletal strain provides an opportunity to study how specific exercises influence strength gains.

Summary of Work Completed

The goals of the C^3RP funded project were to develop detailed human numerical models and predict internal forces during exercises. These goals were met through completion of the work described in this section.

Literature review & software evaluation

Literature reviews were conducted to identify previous work and understand the state of the art in human modeling and internal force prediction. The results of these reviews also facilitated the selection of appropriate software tools for the project.

Prior studies aimed at producing numerical models of individual bones have relied primarily on CT scans. The brightness of a CT scan image is related to the density of the scanned object. In bone, density is related to stiffness. So, a CT scan image can be interpreted to generate the bone geometry and stiffness variation. In most prior studies, the scan data were then processed using custom-designed image processing tools, with significant user interaction. However, a few recent studies use the MIMICS software package to create a 3-D surfaces from the individual scans. This tool was selected for the present project. In addition, after a review of the finite element pre-processor options, TrueGrid was chosen as the best finite element hexahedral mesher.

Prior studies aimed at predicting internal forces during exercises have relied on either a forward- or inverse-dynamics approach. For forward-dynamics, a rigid-body dynamics software package

was often used (e.g. ADAMS or SIMMS). However, reverse-dynamics allows the inclusion of additional data (e.g. muscle activations, through EMG), to improve the prediction capability. The few published inverse-dynamics studies used custom-designed software for predictions. In this project, the SimMechanics module of Simulink was selected for use in implementing an inverse-dynamics analysis.

Obtaining specimens and collecting data

Two specimens were needed for this project:

- A cadaver femur for CT scans and mechanical testing*
- A volunteer to perform instrumented exercises and submit to MRI scans

A fresh-frozen excised cadaver femur was obtained from the University of California – Irvine Medical School's Willed Body Program. A CT scan of the femur was performed at Arroyo Grande Community Hospital. The data from this scan was collected on a CD-ROM. The femur will be tested in bending in the Cal Poly mechanical engineering department's Composites Lab on Tuesday, March 11. Force, deflection, and strain will be collected during these tests, in part using prototype bone strain gages developed by UC-Irvine.

The volunteer selection and data collection was performed according to procedures approved by the Cal Poly Human Subject Committee. The volunteer performed leg extension, leg curl, and leg press exercises while instrumented with surface electromyography (EMG) sensors on the Vastus lateralis, Vastus Medialis, Rectus Femoris, Biceps Femoris, Semitendonosus, and Semimembranosus thigh muscles. Exercises were videotaped for motion data and dynamic force and position data were recorded. Following data collection, the volunteer received two MRI scans at Arroyo Grande Community Hospital, from the pelvis down to the upper calf of the lower limb. The data from these scans was collected on a CD-ROM.

Procedure development

Human Modeling

In order to develop an accurate finite element model from a CT scan, a number of steps must be followed. In prior studies, many of these required extensive user interaction. The intent of the present study was to introduce significant automation to the process. The major steps, and the tools used to complete them in the current study, are:

1. Mask scan data to leave only bone materials (MIMICS – based on CT intensity)
2. Develop internal and external 3-D surfaces for the bone (MIMICS – based on mask)
3. Use surfaces to create hexahedral finite elements (TrueGrid – with computational mesh)
4. Apply material properties to finite elements (MIMICS – with TrueGrid elements)

After obtaining the cadaver femur CT scans, MIMICS was used to mask the bone (Figure 1) and create mathematical surfaces (Figure 2). These steps proceeded with minimal user interaction, saving significant effort over past studies. Unfortunately, at this point MIMICS does not have the capability to generate hexahedral (six-sided element) meshes. So, the surfaces were exported

* A cadaver femur, not part of the original project plan, was selected after doctors expressed concerns with unnecessary CT scans. This additional also enables direct measurement of strains predicted by the developed bone model.

from MIMICS and imported into TrueGrid for processing into a detailed hexahedral mesh of the bone (Figure 3). Unfortunately, processing the femur model in TrueGrid required significant effort to develop an appropriate computational mesh (right) that would map correctly to the surfaces (left). However, now that this computational mesh has been developed for one femur, we believe it can just be scaled and directly used for another femur, reducing user interaction in the future. Similar computational meshes would need to be defined once for each bone type.

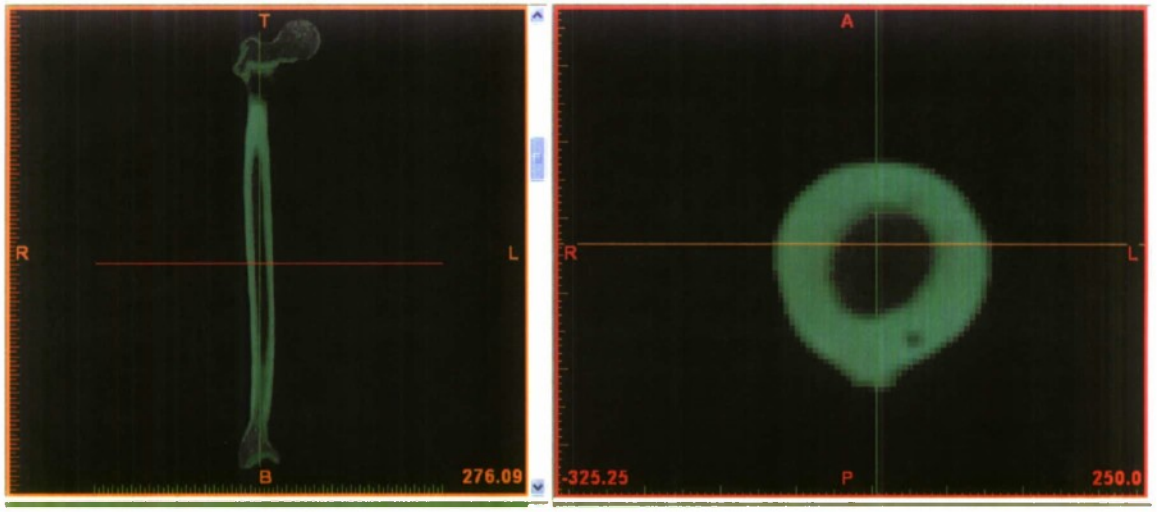


Figure 1: Mask created in MIMICS from the CT scan of the cadaver specimen.

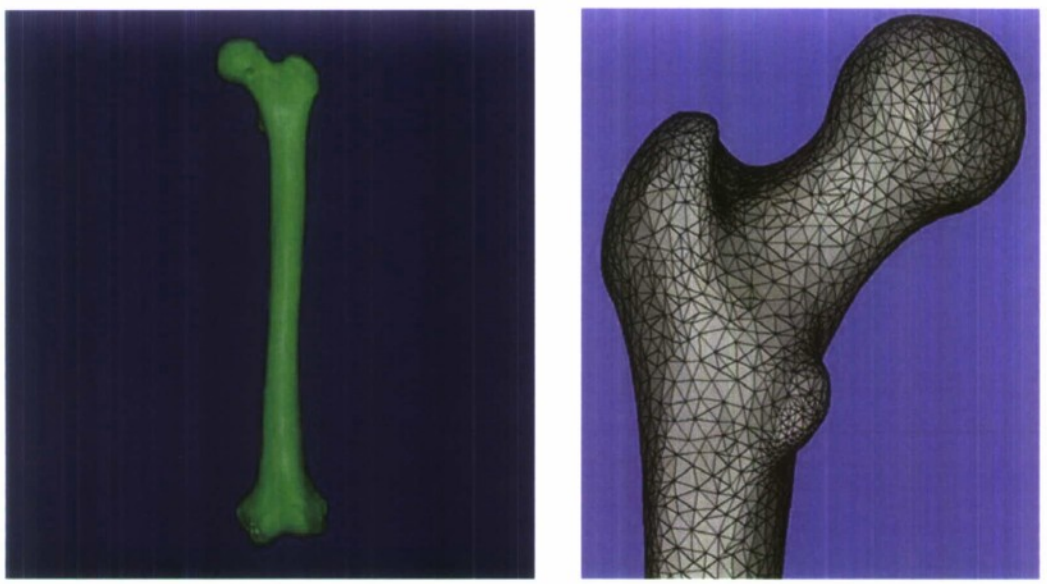


Figure 2: CAD surfaces created in MIMICS from femur CT scan data

The final step in the process, re-importing the model into MIMICS and assigning inhomogeneous material properties, has not been completed. However, similar work has been performed on bovine femurs in the Cal Poly biomedical engineering department, so no issues are anticipated.

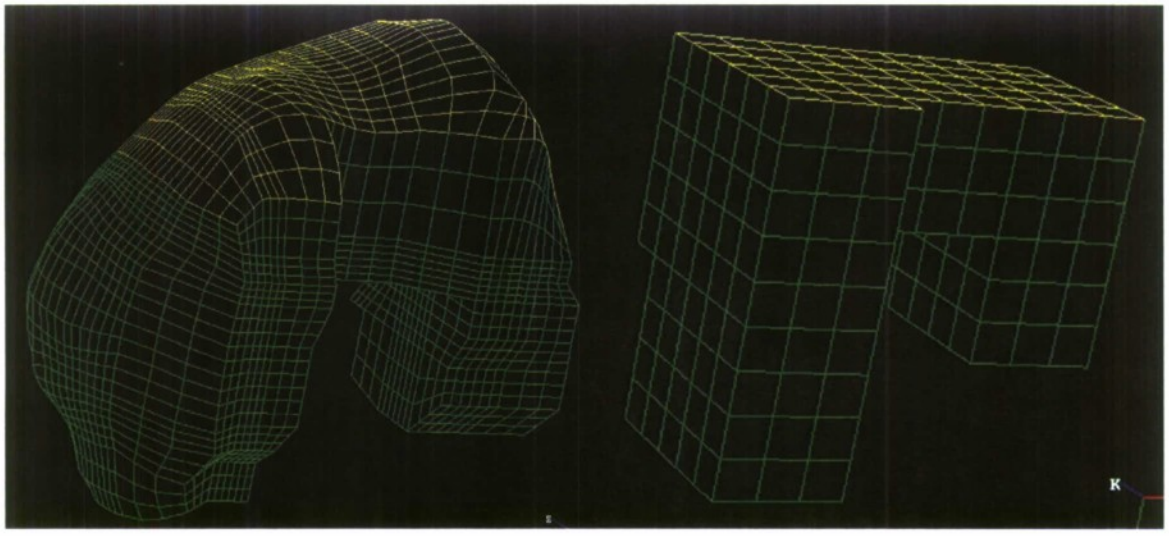


Figure 3: Solid hexahedral mesh (left) developed from a computational mesh (right) in TrueGrid

Internal Force Prediction

In order to develop an accurate prediction of internal (muscle, ligament, and contact) forces during specific exercises using inverse dynamics, a number of steps must be followed. In prior studies, many of these required extensive user interaction. The intent of the present study was to introduce significant automation to the process. The major steps are:

1. Predict surface muscle activation based on EMG data
2. Predict deep and surface muscle forces based on surface muscle activation
3. Develop geometry model of the major interacting bones
4. Assess contact and ligament forces based on known geometry and muscle forces.

Raw EMG data (Figure 4) was normalized using a root-mean-squared (RMS) technique (Figure 5). Muscle activation (%) was found by comparing normalized EMG for a particular exercise to the maximum EMG recorded during a peak isometric exercise. Muscle forces were predicted based on their activation level, current length, and cross-sectional area (measured from MRI scan data - Figure 6). Deep (hidden) muscles were assumed to have the same activation as adjacent surface muscles.

For the current study, a 2-D geometry model was developed to capture forces acting on the femur during to leg extension and leg curl. Most of the forces act in a single plane during these exercises, so this simplification was appropriate. Known muscle lines of action (from the MRI scan) were used to reduce the muscle forces to their in-plane components for this analysis. Contact and ligament forces were combined into a single joint model due to the limited joint alternatives available in SimMechanics. Simplified bone geometry was used in the force prediction model. The SimMechanics block diagram is shown in Figure 7, while the model geometry is shown in Figure 8.

Once the geometry model was created and appropriate predicted muscle forces were added, the joint forces and external reactions (joint torque) were output from the model. The resulting joint torque was used for model validation.

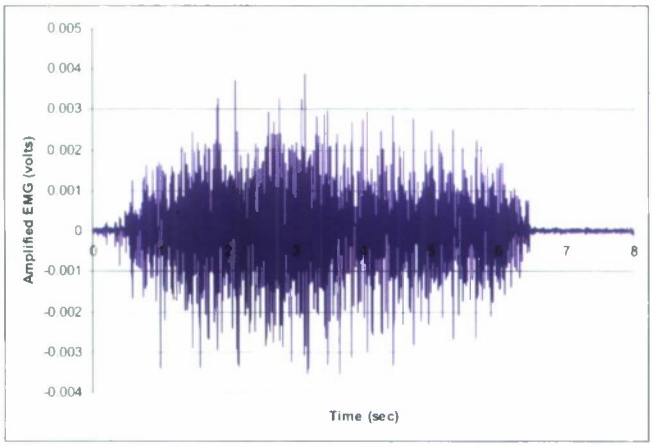


Figure 4: EMG data taken from the Vastus Medialis muscle for the isometric seated knee extension exercise with the subject's leg held at 105°

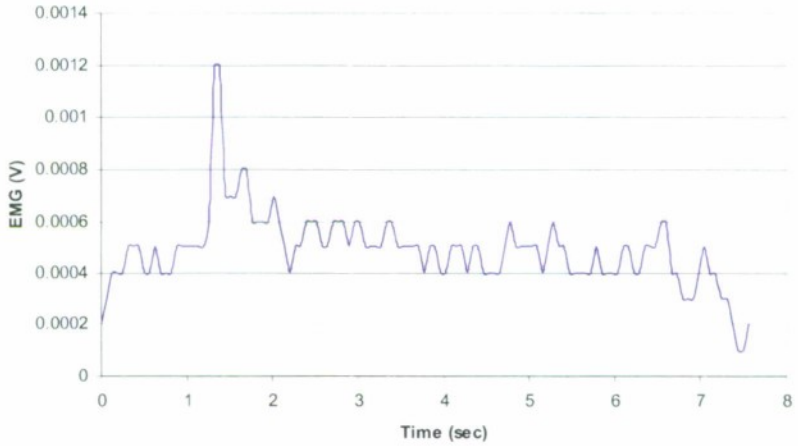


Figure 5. A sample of raw EMG data post processed with the moving window RMS algorithm

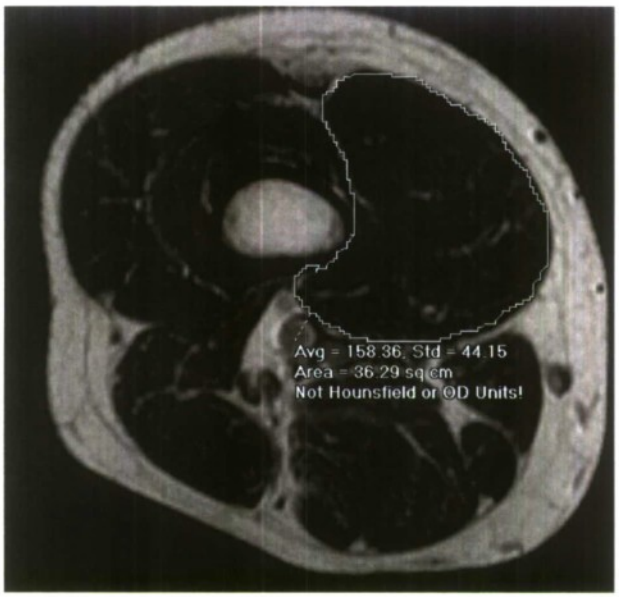


Figure 6. Using subject's axial MRI to determine the Vastus Medialis size

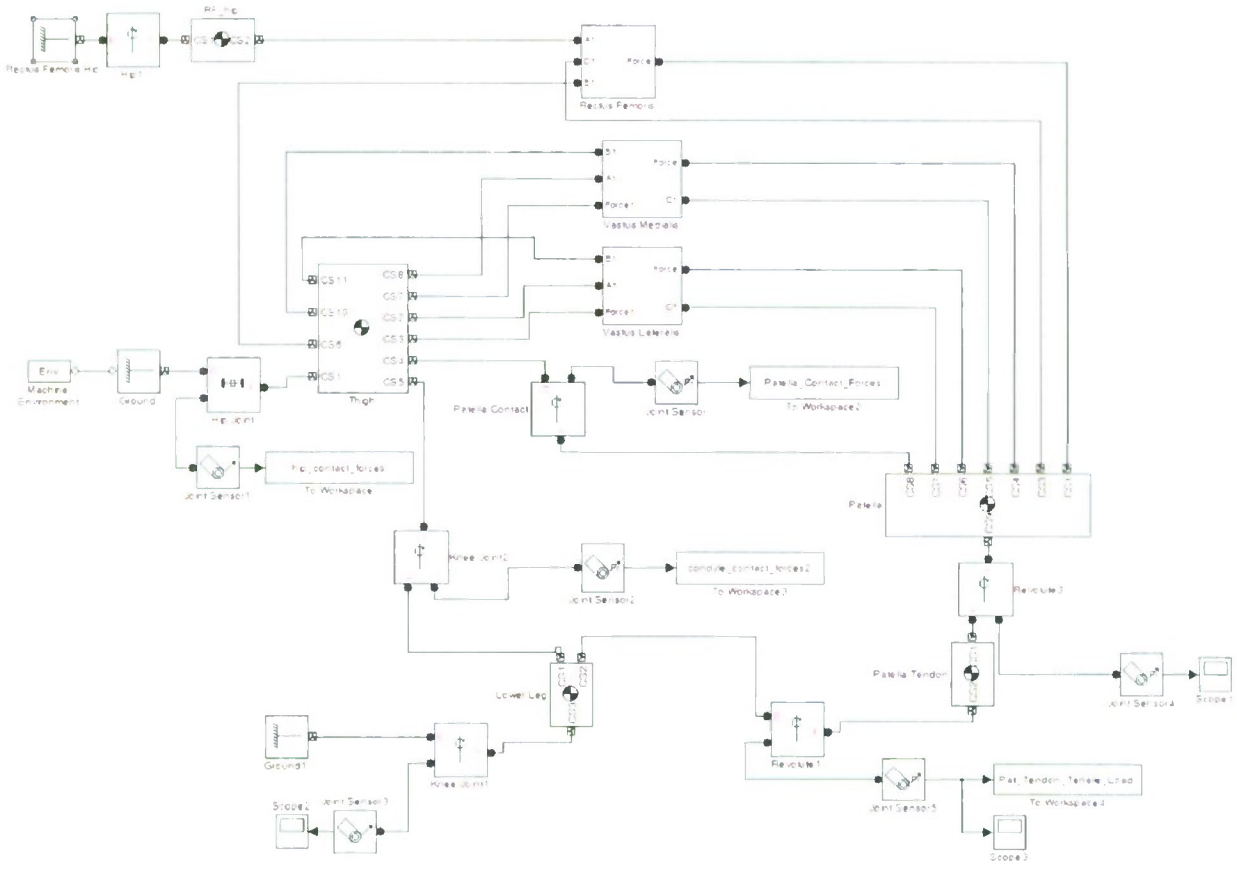


Figure 7: SimMechanics block diagram for internal force prediction

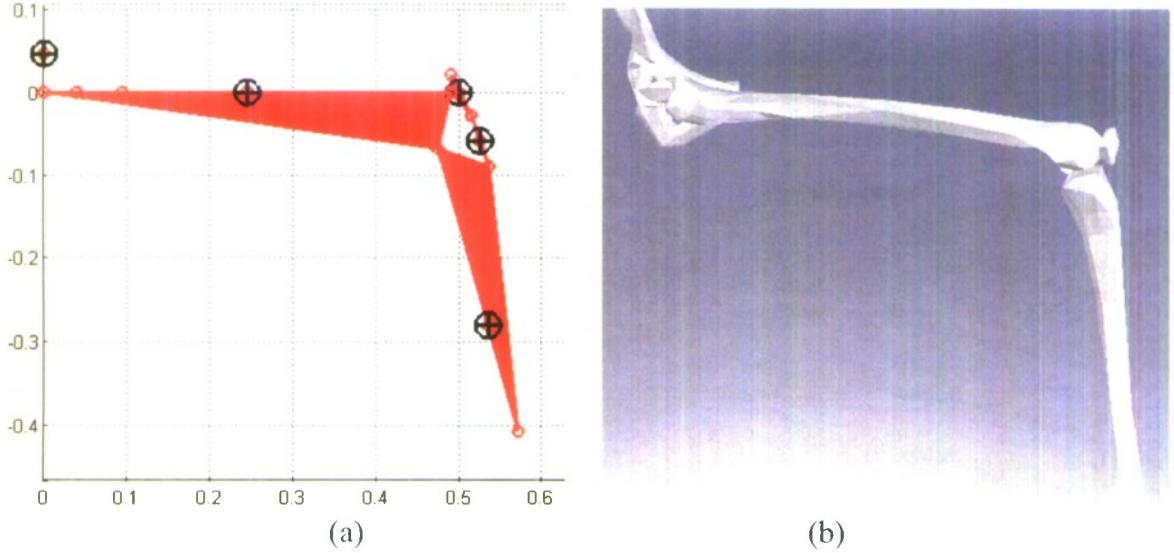


Figure 8: The (a) 2-D SimMechanics model represents the (b) 3-D bone geometry

Procedure evaluation

The success of the human model development procedure will be assessed by how well the resulting bone finite element model ean prediet the loeal strains measured during physieal 3-point bending tests on the eadaver femur. Cadaver femur testing is scheduled for March 11, 2008.

The success of the internal foree prediction procedure will be assessed by comparing the model-predicted output knee joint torque to the measured knee joint torque during the instrumented volunteer exercises. This eomparison will be performed by the end of Mareh, 2008).

Collaboration

The development of the procedures has opened up on-going interactions with academic partners at Cal Poly (Scott Hazelwood in Biomedieal Engineering and Bob Clark in Kinesiology) and UC Irvine (Joyce Keyak), which should lead to collaboration on future projeets. Several of the challenges eneountered during the proeedure development have identified future researeh directions for these collaborations.

Multi-Disciplinary Project Based Learning

One of the objeetives of this project was to foster learning and researeh at Cal Poly. This was accomplished by involving students from different levels and backgrounds. Two mechanical engineering graduate students, Estevan Negrete and Matthew Seott, formed the eenter of the student team by leading the development of the two proeedures. In addition, one mechanical engineering and two kinesiology graduate students partieipated in the projeet. As a result, this project has contributed direetly to the edueation of five Cal Poly students.

Conclusion

Two major steps have been completed toward understanding the link between specific physical exereises and resulting local bone strength changes. In the present projeet, a detailed orthotropie finite element model was developed for a human femur, based on CT sean data. The model will be cheeked for validity to 3-point bending test results of the same femur. In parallel, data collected during exereises with a human volunteer were used to prediet the internal forees acting on the bones, using a reverse-dynamies model. Validation of both models will be completed in the next month.

References

Cussler EC, Going SB, Houtkooper LB, Stanford VA, Blew RM, Flint-Wagner HG, Metcalfe LL, Choi J, Lohman TG, "Exercise frequency and calcium intake predict 4-year bone changes in postmenopausal women." *Osteoporosis International* **16**(12): 2129-2141, 2005.

Engelke K, Kemmler W, Lauber D, Beeskow C, Pintag R, Kalender WA, "Exercise maintains bone density at spine and hip EFOPS: a 3-year longitudinal study in early postmenopausal women." *Osteoporosis International* **17**(1): 133-42, 2006.

Kato T, Terashima T, Yamashita T, Hatanaka Y, Honda A, Umemura Y, "Effect of low-repetition jump training on bone mineral density in young women." *Scandinavian Journal of Medicine & Science in Sports* **16**, 2006.

Kemmler W, Weineck J, Kalender WA, Engelke K, "The effect of habitual physical activity, non-athletic exercise, muscle strength, and VO2max on bone mineral density is rather low in early postmenopausal osteopenic women." *J Musculoskelet Neuronal Interact* **4**(3): 325-34, 2004.

Lorentzon M, Mellström D, Ohlsson C, "Association of amount of physical activity with cortical bone size and trabecular volumetric BMD in young adult men: the GOOD study." *Journal of Bone Mineral Research* **20**(11): 1936-43, 2005.

Miller LE, "Bone mineral density in postmenopausal women." *Physician & Sportsmedicine* **32**(2): 18, 2004.

TrueGrid, April 2006. TrueGrid User's Manual. Compiled by Rainsberger R, XYZ Scientific Applications, Inc. 1992 – 2006.

U.S. Department of Health and Human Services (DHHS), *Bone Health and Osteoporosis: A Report of the Surgeon General*. http://www.surgeongeneral.gov/library/bonehealth/docs/full_report.pdf. October, 2004.

Vainionpää A, Korpelainen R, Vihriälä E, Rinta-Paavola A, Leppäluoto J, Jämsä T, "Intensity of exercise is associated with bone density change in premenopausal women." *Osteoporosis International* **17**(3): 455-63, 2006.

Solar Transportation: Sunlight to Electricity to Motion

Project Investigator:

Peter V. Schwartz,
Cal Poly Physics Department
California Polytechnic State University
San Luis Obispo, CA

C^3RP funding was used to initiate our present studies in renewable energy. The work can be broken down into three parts: Solar Concentrators, Electric Vehicle Simulation, and Analysis of Energy Substitution.

1) Development of a novel way to inexpensively generate solar electricity with a nontracking solar concentrating mirror assembly. The fact that the main concentrating mirror doesn't move would allow construction of large surface areas of solar concentrators by simply digging the shape into the earth and covering the surface with a reflective coating. This method of solar electricity production may be ideal for the military. Solar troughs could be rapidly deployed upon arriving at a new location, and disassembled before leaving.

We are developing a novel solar concentrating technology that does not require the primary focusing element to be rotated. If sunlight is focused onto an absorbing element with a stationary fixture, the focused light will move as the angle of the sunlight changes (Fig.1). Our particular efforts are dedicated to a low cost, lightweight, easily deployable method to focus sunlight in the field. The focused sunlight can be directly turned to electricity with a high efficiency photovoltaic (HEPV), or used to heat a working fluid for a microturbine (Solar Thermal Electric, STE), that could provide both heat and power.

Both students have made great progress in using ray-tracing software in order to simulate solar concentrators (Fig.2). The design is far from optimized, but shows significant efficiency and solar concentration. Additionally, we have built a primary mirror for our first prototype (Fig.3).

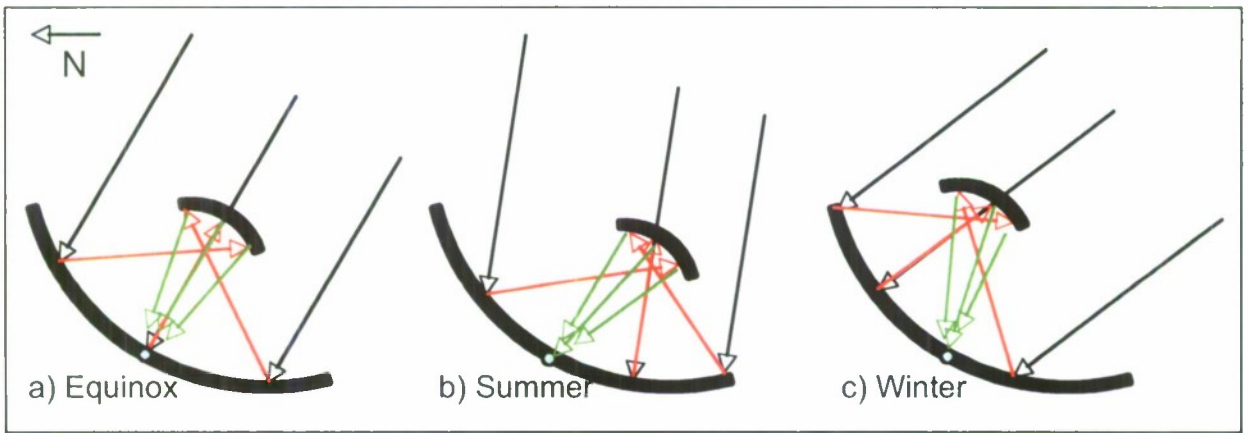
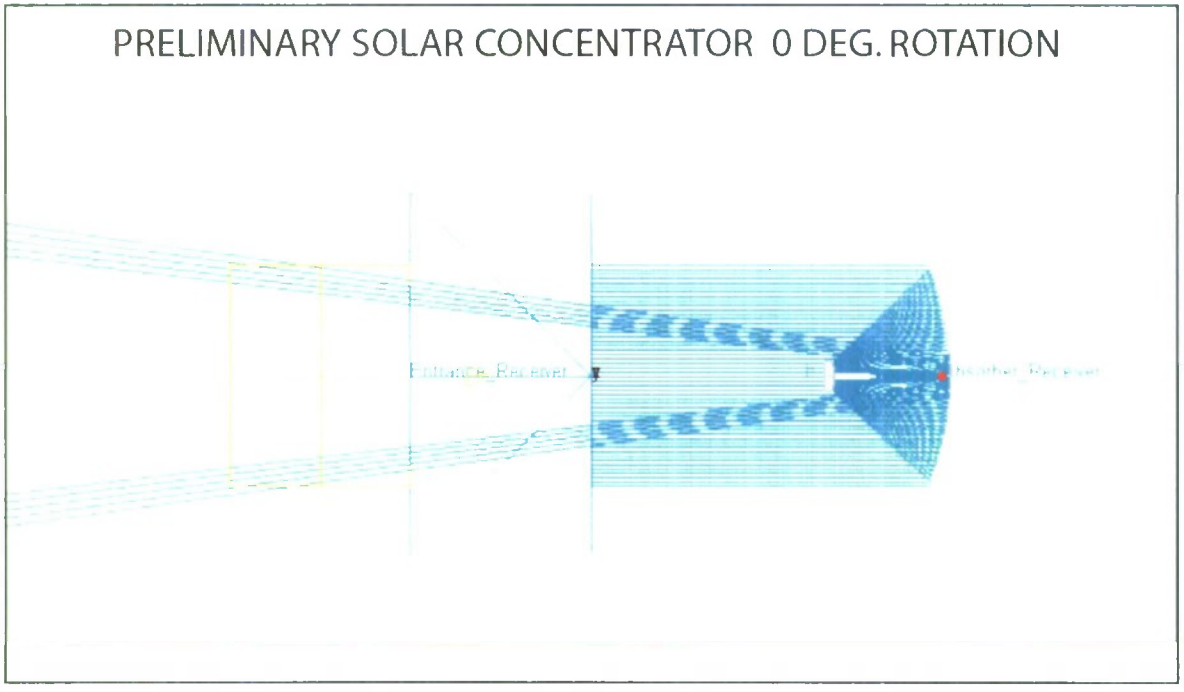
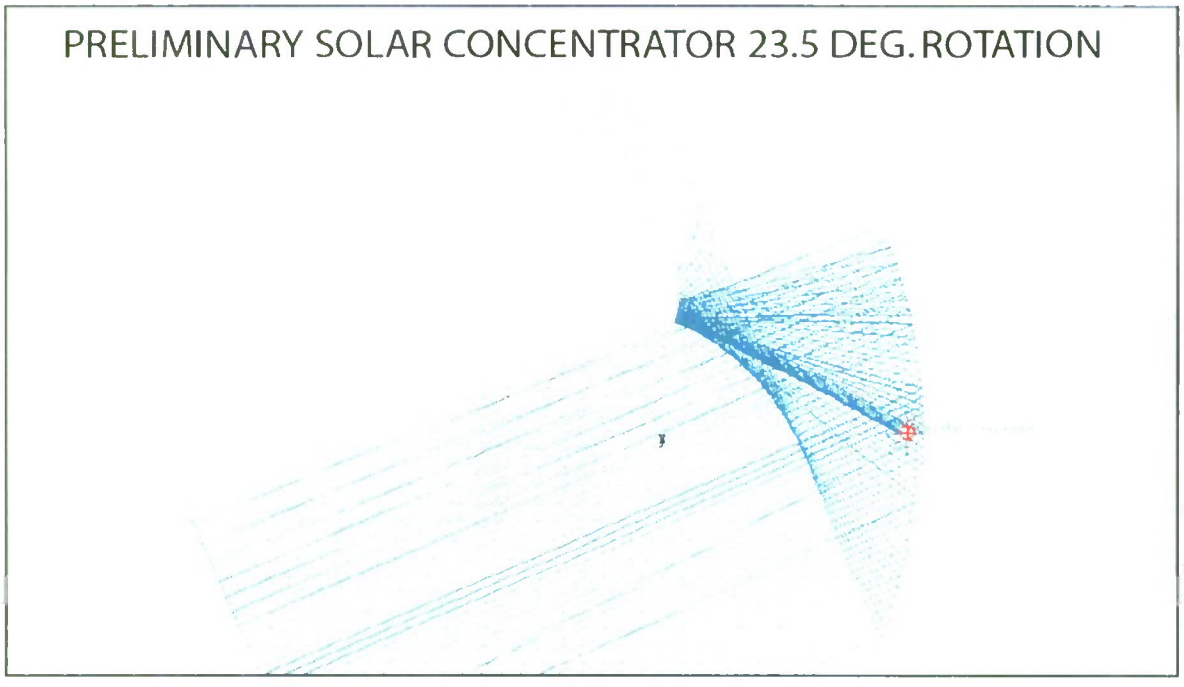


Fig.1 The secondary mirror can be located beyond the focus of the primary mirror to optimize performance. Black rays represent incident light, red rays represent light reflected from the primary mirror, and green rays represent light reflected from the secondary mirror. If the surface is optimized for equinox, summer and winter light is still focused well onto the same area.





Geometry (4ft)	Absorber Surface Area ASA (m2)	Entrance Aperture EA (m2)	Light Source Angle (deg)	Number of Rays Thru Entrance (N)	Number of Rays Thru Absorber (N')	Power Transmission Factor (N'/N)	Geometric Concentration Factor (EA/ASA)	Flux Thru Absorber (W/m2)	Sun Power (Suns)
0Parabolic	0.06	1.05	0.00	10000	6387	0.64	17.61	9227.84	11.25
0Parabolic	0.06	0.94	23.50	10000	3879	0.39	15.62	4555.38	6.06

Fig 2 Ray tracing diagrams of a preliminary reflecting geometry. Initial concentrations are between 6 and 11 suns. While this is a start, further modeling with be dedicated to increasing concentration and minimizing lost rays.

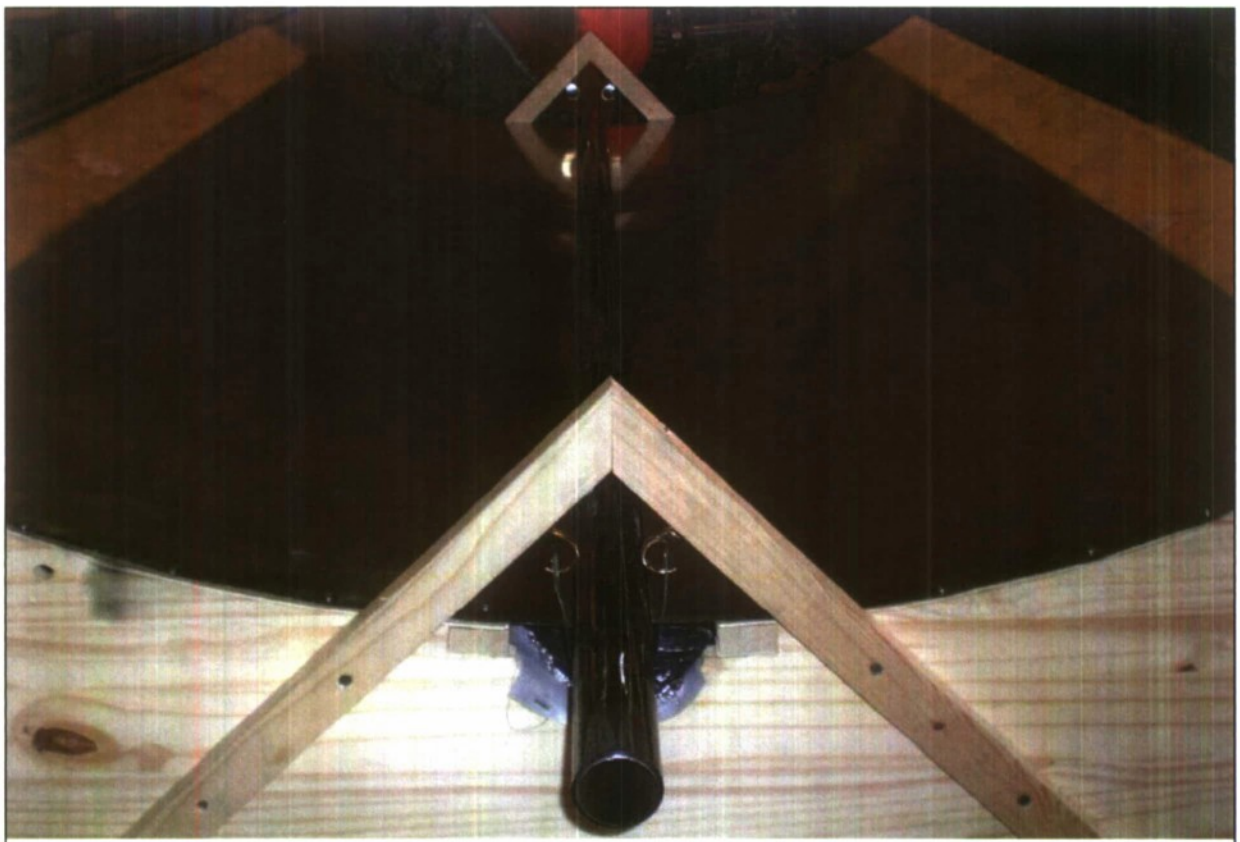


Fig 3. Primary mirror and Light Absorbing Element. The design allows for movement of the light absorbing element.

2) Development of a virtual electric car in order to study battery technology. With data from battery and automobile manufacturers, we are designing a driving simulator that will predict energy use and performance in a hypothetical electric vehicle.

Three Physics Students (C^3RP funded one of the three students) have been studying and modeling the charging and discharging behavior of the new, high energy density, fast charging Lithium Ion batteries. We have produced a simulation of the batteries and the electric motor and controller as well as the physical parameters of driving such as friction, mass, and desired acceleration.

3) Analyzing technology substitution. When substituting renewable electricity for gasoline in an automobile, or for natural gas to heat water it is important to understand the full "life cycle" costs: How much does it cost financially? What are the energy conversion efficiencies? How much carbon dioxide is emitted for each mile traveled, and how much land is used for each mile traveled?

While the overarching challenge we address is how to electrify the transportation sector, this proposal is particularly dedicated to the production of electricity and heat through solar concentration.

Renewable energy and increased efficiency can be substituted for conventional energy use. We have been analyzing this substitution with regard to financial costs, environmental costs, and feasibility. Over the summer, two students improved an economic energy model for Santa Barbara County that simulated a transition to fossil neutrality. Much of the work involves analysis of transportation energy use. One project compares energy flows for biofuels (such as corn ethanol) and solar electricity. The project was picked up last fall by two physics majors and is progressing well. I have given several talks on the results of this model on campus and for the public in San Luis Obispo. C^3RP funding supported one student this summer.

Use of Ultrasonic Sensors in the Development of an Electronic Travel Aid

Project Investigators:

Brian Self, Chris Gearhart, Alex Herold,
Charles Birdsong and Lynne Slivovsky
Mechanical and Electrical Engineering
California Polytechnic State University
San Luis Obispo, CA

Abstract

Ultrasonic sensors present one of the most cost-effective digital distance measurement systems available for mobile applications. Their effectiveness is limited, however, in applications involving complex environments and when information on sensor position is unavailable. This paper focuses on the implementation of ultrasonic sensors and associated considerations during development of an Electronic Travel Aid [ETA] for the visually impaired utilizing ultrasonic sensors. These sensors will later be interfaced to a vibrotactile feedback system that will provide navigation cues to the wearer. Our work with sensors included signal filtering and triangulation to improve performance characteristics of ultrasonic-based measurements. We also describe the use of computer modeling to aid in the design of ultrasonic sensor systems. Finally, preliminary testing with a vibrotactile belt system will be discussed.

1. Introduction

The creation of a viable electronic travel aide for people with visual impairments has been the topic of substantial research for many years, with different efforts achieving varying degrees of success. Electronic travel aids have expanded to include the use of global positioning system technology, wireless network communication, ultrasonic sensors, and computer vision systems. [2][3][4] As mobile computing technology has improved, ETA systems have grown more complex. This system exclusively utilizes ultrasonic sensors combined with vibrotactile feedback in order to minimize the complexity of the system for the end-user.

Many existing ETAs are not well accepted by the visually impaired community. This is not to say there is any deficiency in their design or functionality, simply an observation that although many systems are available, none has become as common as a white cane. When considering the significant improvements that have taken place with ETA technology and the long track record of research in this field, the slow adoption of these systems suggests that existing designs do not meet the users' requirements. We concluded, in part, that one reason for the slow adoption rate is the cost of such systems, which have not yet reached a point of economic viability for an average consumer. In order to move closer to economic viability, the cost of the end product was a significant consideration.

Furthermore, as the capabilities of ETAs increase, so does the complexity of use. People with visual impairment must devote significantly more effort when navigating unfamiliar environments than do sighted people. They are unlikely to use a system if it distracts them from the skills they must acquire in order to be comfortable with nothing more than a white cane. We concluded that any system should minimize the potential to distract users.

To that end, we selected tactile feedback to communicate with users. [7] Tactile feedback utilizes vibrating mechanisms in contact with the body to present information. Perhaps the most ubiquitous example of tactile feedback in current technology is the "vibrate" mode on a cell phone. Tactile feedback was selected over auditory feedback because people with visual impairment typically utilize sounds from the environment for navigation, and auditory feedback has a greater potential to distract users. [4]

This paper focuses on our use of ultrasonic sensors in an electronic travel aid, and discusses the difficulties associated with using ultrasonic measurements in unpredictable environments. These include wide variation of target object properties and measurements taken without tracking sensor source location. Tracking sensor location can improve functionality [1], but increases cost and complexity because of additional sensors required to measure movement.

2. Triangulation

Triangulation is the application of geometric relationships to scalar measurements to calculate the position vector of an object. In general, triangulation requires at least the same number of sensors as the number of dimensions within which to fix the position of an object – two sensors can identify location in a 2-D plane, while three sensors can fix location in 3-D space. Additional sensors can be added to improve resolution and accuracy. [1]

We chose to use two sensors because the tactile display we selected can provide feedback on lateral position and distance – not height. Figure 1 shows a scenario that typically results from using a single ultrasonic sensor. Objects at any point on the arc (e.g. point A or B) yield the same reading when measured by a sensor at S_1. Without additional information the

system cannot tell the user how to avoid an object, only that an object is present in the field.

Figure 1 – Point A and B have the same scalar distance from a single ultrasonic sensor at S_1 and cannot be distinguished.

Figure 2 demonstrates how the addition of a second sensor fixes object location in two dimensions. While readings from the left sensor (blue arrows) are the same as in Figure 1, readings from the right sensor (red arrows) have different lengths (as evideneed by the red arcs passing through eaeh point) – creating distinction between point A and point B. This allows our system to prompt the user to move left to avoid an object at B, or right to avoid an objeet at A.

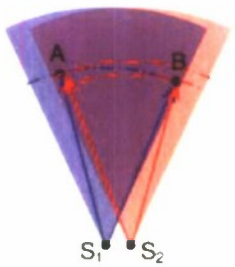


Figure 2 – Triangulation from S_1 and S_2 creates distinction between points A and B.

Tracking sensor movement between readings allows triangulation to be used in conjunction with a voting method in order to approximate an environmental map for path planning. [1] Reeording sensor movement is neeessary to determine whether the same object is detected in sueeessive readings. We opted not to measure sensor movement in order to reduce the number of components in our system; instead, using triangulation to provide more preeise navigational cues.

2.1 Equations for Magnitude & Angle

Our system utilizes measurements taken from two sensors at distinet locations to ealeulate the position vector to an objeet relative to any point along a line connecting both sensors. Figure 3 shows the naming convention used throughout this paper for geometric parameters used in triangulation. Based on the length separating the sensors, L, and readings R_1 and R_2, we triangulate the distance D and angle θ between points O and P. The angle α_n is the angular position of sensor n. The angle γ is half of the angle from the detection beam centerline to the edge of the detection area, and is constant for a specific sensor. Finally, β_n is the angle to a specific reading within the detection field relative to sensor n.

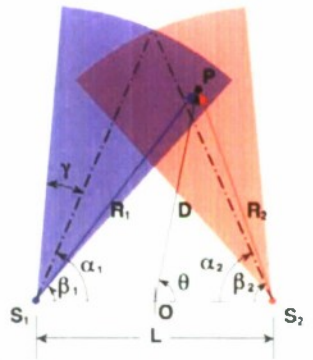


Figure 3 – Diagram of geometric parameter naming convention used to derive triangulation equations.

To calculate D and θ, we derived equations by applying the law of cosines to the system in Figure 3. These equations assume the following: point P exists within range of both sensors, two-dimensional triangulation, sensor one (S_1) is polled first and given priority in all subsequent calculations, L is constant, and since only the positive value of D is physically valid, the negative value given by the square root is intentionally omitted. Equations 1-3 are consistent with those established by Wijk, et al [1], differing only because they take vector form rather than Cartesian coordinates.

$$\cos\beta_1 = \frac{L^2 + R_1^2 - R_2^2}{2LR_1} \quad (1)$$

$$D = \sqrt{R_1^2 + L^2/4 - R_1 L\cos\beta_1} \quad (2)$$

$$\theta = \cos^{-1}\left(\frac{R_1^2 - L^2/4 - D^2}{LD}\right) \quad (3)$$

2.2 Valid Triangulation Limits

When ultrasonic sensors are used in complex environments, scenarios may occur in which equations 1-3 have mathematically valid solutions that have no physical signifieanee. Figure 4 shows two sensors deteeting separate objeets. The objeet at P_1 is detected

from S_1, and the object at P_2 is detected from S_2. Applying equations 1-3 to this scenario will result in triangulation of point P_T. P_T is a mathematically valid solution to the equations, but it is not located within the region we can triangulate (the overlap between the shaded areas).

In these cases, the navigational instructions given by the system would not assist in collision avoidance and only interfere with processing valid data. In order to prevent erroneous data from corrupting valid data during processing, we derived limits for triangulation results based on the characteristics of sensor detection region geometry as criteria for rejecting triangulation results.

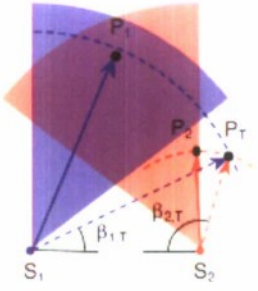


Figure 4 – Typical triangulation error associated with multiple objects in the detection region.

To prevent the error demonstrated above, we observe that angles $\beta_{1,T}$ and $\beta_{2,T}$ are outside the valid triangulation region (the overlap of shaded regions in Figure 4). Referring back to Figure 3, the left sensor detection region is bounded by angle $\beta_1 = \alpha_1 \pm \gamma$. Thus, the valid range of equation 1 is given by:

$$\cos(\alpha_1 - \gamma) \le \cos\beta_1 \le \cos(\alpha_1 + \gamma) \qquad (4)$$

Results from the left sensor are governed by equation 1, subject to the limit in equation 4. In order to establish similar boundary conditions for the right sensor, we first need an equation describing the right sensor similar to equation 1. Solving the law of cosines for the angle β_2 yields:

$$\cos\beta_2 = \frac{L^2 + R_2^2 - R_1^2}{2LR_2} \qquad (5)$$

If we assume that $\alpha_1 = \alpha_2$, then the system simplifies because, by symmetry, the limits in equation 4 apply to equation 5 as well. Furthermore, angles α_1, α_2, and γ are constants, and equation 4 can be pre-computed to obtain constant minimum and maximum values for equations 1 and 5.

We utilized these simplifications to optimize our algorithm for application in a microcontroller. We included the constant values directly in the source code of our program to reduce processing time. Additionally, equations 1 and 5 do not calculate the angles β_1 and β_2 because only the cosines of these angles are required to find θ; this avoids calls to trigonometric functions except for finding the angle θ.

As a final note on these limits, consider Figure 5 below, where the model specified by the sensor manufacturer is compared to the cone approximation we used to develop detection limits. This demonstrates that the limits we derived are conservative values relative to the outer edges of the detection region, and liberal with respect to long-range detection. If desired, the cone approximation can be intentionally shortened (shown as the solid black line) compared to actual detection in order to make all limits conservative.

Figure 5 – Comparison of sensor manufacturer detection area with cone approximation.

3. Sensor System Modeling

As part of the design process, we modeled two sensors in MATLAB to optimize the sensor configurations. This section outlines the development of our model and compares the results of the actual system to the predictions of the model. We tested sensor arrangements after simulation by comparing direct measurements of object position taken with a tape measure with ultrasonic readings. Other systems compare readings (e.g. calculated environmental map) with measurements of the real environment when the system is in use, which tests the performance of the system. [1]

System simulation is beneficial because it can be performed prior to construction to generate system design parameters. This is helpful because it can be done based on the manufacturer's published specifications, allowing a wide range of ultrasonic products to be evaluated based on resolution, range, and detection area (among other factors) for their performance benefits.

While simulation is helpful during the design and testing of a system, it is important to recognize that modeling is not a substitute for testing and validation.

The results obtained by modeling must be verified after construction to ensure that the characteristics for which the sensors were chosen actually provide the desired effect.

The simplest model consists of a plot of manufacturer specified sensor detection area adjusted for specific configuration for a single instant in time. Such a model assumes the effects of object orientation with respect to the sensors, material, and size, are negligible. These are valid assumptions if the type of object to be detected will remain uniform throughout the operation of the system.

We focused on triangulating medium size objects, defined by the manufacturer as 3.25" diameter rods, or equivalent. To that end, we used our model to find a sensor orientation that maximizes system performance for triangulation of medium size objects. We were specifically interested in the rotation angle of the sensors (angle α from Figure 3) for a fixed separation length, L, to achieve "best" detection.

We defined "best" detection based on two constraints: maximizing the average width of the triangulation area (which provides the greatest range of available angles during triangulation), and maximizing the number of intersections of ares based on the resolution of the sensors (which corresponds to the number of points where triangulation can occur). Finding an orientation that maximizes both constraints provides the largest usable triangulation region.

The maximum average width of the detection area occurs when both sensor detection areas overlap completely (i.e., no separation between the sensors and no rotation). Except for this case, the overlap region will always be a subset of, and smaller than, the total detection region of a single sensor. As a result, this constraint draws the sensors close together.

The average width of the overlap region is a meaningful criterion because it provides a measure of triangulation effectiveness for navigational instructions to avoid obstacles. When the width is small, as in Figure 6(a), triangulation cannot take place in the regions scanned by only a single sensor – limiting the instructions that can help avoid those objects. A greater width, shown in Figure 6(b), allows triangulation to proceed for the majority of the detection area.

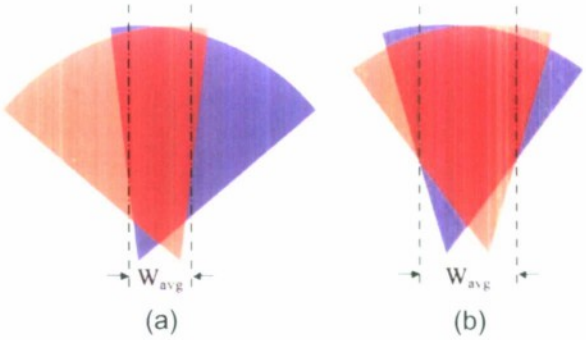


Figure 6 – **(a)** Thin overlap region, most detection space lies outside of triangulation area. **(b)** Wide overlap region, most detection area is inside triangulation region.

Ultrasonic sensors report the distance to objects to the nearest unit of resolution (frequently the nearest inch or centimeter). By drawing arcs at each unit length in the detection cone and finding the intersections of ares from both sensors, it is possible to map all valid triangles for sensors in a specific orientation. Our second criterion for effectiveness counts the number of arc intersections – which draws the sensors apart and introduces rotation, as shown in Figure 7 (a) and (b).

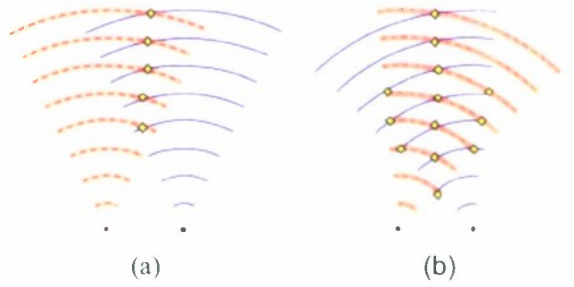


Figure 7 – Triangulation is possible at every point where two arcs intersect, represented here as yellow dots. **(a)** Separating the sensors creates intersections. **(b)** Rotating the sensors can increase the number of intersections, creating more opportunities for triangulation in the same area.

3.1 Binary Image Modeling of Ultrasonic Sensor Detection Areas

This section describes the method we used to model system orientations with images. Our model was designed to find the effects of sensor orientation on the effectiveness of triangulation, and to estimate the change in effectiveness caused by small changes in sensor orientation. We were concerned users might

move the sensors for better comfort, or that the sensors might shift from their original position during normal user movement.

Our model used drawings of detection regions from the sensor manufacturer and image analysis methods to approximate sensor performance. By using images, arbitrary shapes can be used to draw the detection area – eliminating the need for a precise mathematieal model. This allows the model to be changed quickly, rather than developing new equations.

The theoretical detection area, found in the datasheet of the Maxbotix LV-EZ1 sensor, is capable of detecting a 3.5-ineh eylinder at 5V of power (see Figure 8(a)). It shows detection to a maximum range of approximately 10.5 feet for this size object. The image was imported in Matlab as a binary image and cropped to remove empty spaee from all edges.

Figure 8(b) shows the detection region rotated and corrected for the separation length between sensors. The correction is accomplished by padding the right edge of the image with columns filled with zeros equal to the number of columns separating the sensors.

Finally, Figure 8(e) is a composite of Figure 8(b) to a copy flipped about a vertieal axis. The normalized composite image displays the overlap of both sensor deteetion areas as white, while the area scanned by only one sensor is light grey.

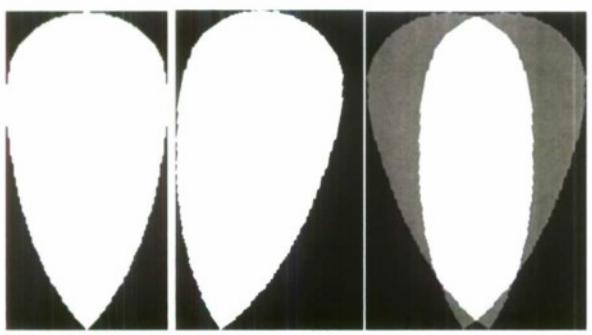


Figure 8 (a) Model of detection area. **(b)** Model image rotated and adjusted for sensor separation length. **(c)** Composite image (normalized) of left and right side sensors.

We calculated the average deteetion area width by determining the average width of the white region in Figure 8(c). To do this, we eliminated all pixels with a normalized value less than one, counted the number of non-zero pixels in each row, and averaged the results.

In order to count the number of possible triangulation points, we fragmented the detection area image from Figure 8(a) into arcs spaeed at the resolution of the sensors, whieh was one ineh. The arcs must be greater than one pixel in width, and separated by a gap of more than one pixel. These conditions are used to ensure that eight-pixel connectivity can be assumed when counting connected clusters. If we assumed four-pixel connectivity, the same cluster might be counted multiple times if it is connected along a diagonal. Separating the arcs by more than one pixel ensures that there will be no diagonal connections in the composite image shown in Figure 9(b) below.

The example in Figure 9 simplifies and exaggerates our method to count triangulation points. Figure 9(a) is an individual sensor image separated into arcs and Figure 9(b) is a eomposite image of two sensors separated an arbitrary distance. Each cluster in Figure 9(c) represents an intersection of two arcs – by counting the elusters, the number of intersections can be approximated.

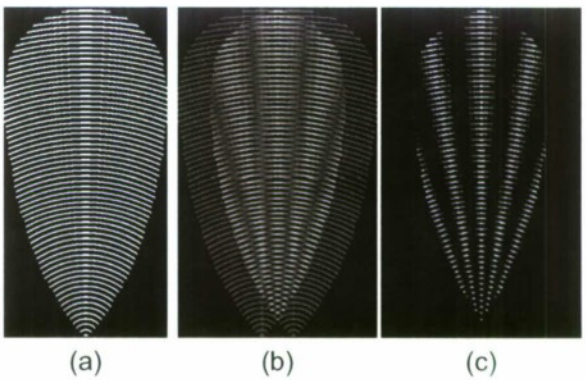


Figure 9 (a) Sensor area represented as separate arcs. **(b)** Composite image of overlap for two sensors. **(c)** Regions from (b) approximating arc intersections.

Our simulation has two limitations. First, it does not aecount for the effects of large sensor rotation angles on aeeurate object detection. We did not aeeount for the effects of large rotation angles because the detection region image is a conservative estimate. Second, the detection area image from Figure 8(a) is only valid for a specifie size of objeets. To overcome the seeond limitation, we repeated the simulation with images valid for other object sizes.

3.2 Results of Modeling

In order to reach general conelusions about sensor orientation we ran the simulation over a variety of distance and angle combinations. This section outlines the results of our simulation for a detection region eapable of detecting a 3.5-inch diameter eylinder.

When designing our system, we decided to mount one sensor on each of the user's shoulders, thereby limiting the possible separation length between sensors. For the initial simulation, we bounded length at less than twenty inches, and rotation angle at less than eight degrees.

Figure 10 and 11 show the results of the study with average width and intersection count plotted as a function of the two parameters length and angle. The results indicate that the average width of the sensor detection region is greatest at small sensor separation lengths, and decreases as either the rotation angle or separation length increase. Furthermore, the number of intersections does not vary significantly based on rotation angle for small separation lengths, approximately ten inches or less. However, the effect of rotation angle on the number of intersections becomes significant at larger separation lengths.

The simulation results suggest that changes in separation length and rotation angle increase the number of intersection points more rapidly than the same changes decrease the average detection area width. We concluded that triangulation with two sensors is most effective for separation lengths greater than twelve inches with rotation angles greater than five degrees.

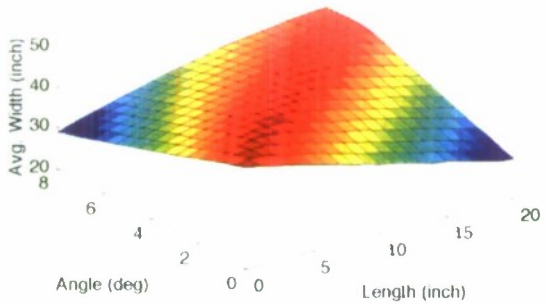


Figure 10 – Variation of average detection area width over allowable separation lengths and rotation angles for two ultrasonic sensors.

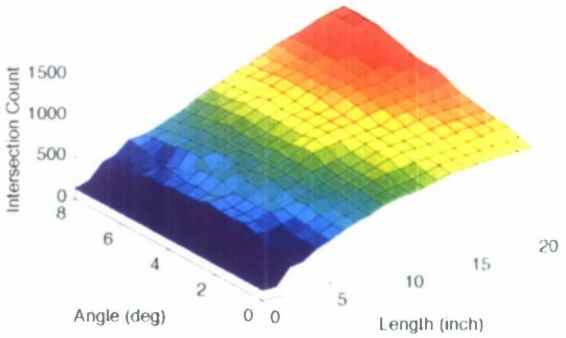


Figure 11 – Number of intersections, representing possible detection points, as separation distance and rotation angle vary.

3.3 System Testing

We tested two sensors spaced ten inches apart and angled towards each other five degrees from parallel in order to test the accuracy of the simulation. We placed a 3.5-inch diameter cylinder at several points near the edges of the theoretical detection range and took readings. After each reading, we moved the object towards the edge of the detection region to find the outer boundaries.

Each red dot in Figure 12 is located at the furthest position at which ultrasonic measurements correctly triangulated the stationary object. We mapped these points over an image of the theoretical detection area to see how well they agree. The results obtained show that the actual detection area closely corresponds to the predicted area.

Figure 12 – Comparison of furthest successful test measurements to theoretical detection area; red dots represent ultrasonic measurements superimposed on the theoretical detection area.

4. Filtering Ultrasonic Sensor Readings

Filtering the raw ultrasonic signals is an important part of the process. Fluctuations due to rapidly changing scenes or spurious electronic signals can cause the system to give misleading information to the user. In this section we describe the different filtering methods that were applied to the ultrasonic distance signals. We incorporated a sliding window filter with either a Kalman filter or a Gaussian low-pass filter to stabilize the signal.

Other systems contain similar filters intrinsically; the primary distinction is that we treat the readings as a continuous signal and apply the filter separately. The voting method used by Wijk et al., [1] is essentially a threshold filter based on the number of votes received. The potential field sensor, developed by Veelaert and Bogaerts [3] doesn't collect distance measurements at all, but constructs the potential field directly from the ultrasonic response. Constructing the potential field creates a buffer that tolerates errant readings because the system does not need to track objects moving through the field.

4.1 Sliding Window Filter

A sliding window filter is a threshold filter, similar to a median filter [5] used in computer imaging. This filter corrects for instabilities when detecting small objects by assuming that there is a threshold that describes the maximum distance an object can move between consecutive readings. The purpose of the sliding window is to prevent a failure to detect small objects in every reading from corrupting the operation of the Kalman or Gaussian filter.

When the sensors miss an object, they return the maximum possible value. The filter compensates by searching a set of previous values to determine if a recently detected object has been lost. If there is a non-maximal value within the window, then the filter assumes that the sensor missed an object. In this case, the filter returns the most recent non-maximum value within the window as the current position, which assumes the true object position is near the most recent previously measured position.

4.2 Kalman Filter

We added a Kalman filter [6] to improve the response characteristics of the sliding window filter. The sliding window is insufficient on its own because the output becomes static when successive sensor measurements fluctuate. Kalman filters use a kinematic model of motion to approximate physical system behavior. When sensors detect objects in rapid motion, they are assumed to continue their path of motion instead of assuming they remain stationary during a sequence of unstable measurements.

4.3 Low-Pass Filter

We also implemented a Gaussian low-pass filter as an alternative to the Kalman filter. A Gaussian low-pass filter assumes signal noise is normally distributed and random. The size of the filter mask (i.e., the number of data points considered) is dependent on the assumed standard deviation. Gaussian filters require data points both before and after the filtered reading to operate. This creates a lag in response between the current measurement and the filter response. The lag induced by the Gaussian filter highlights an advantage of the Kalman filter, since the Kalman filter only requires one previous data point after it is initialized to calculate the current point.

4.4 Comparison of Filter Response

The sensors frequently miss small objects during normal operation, therefore we decided to test the filters on this difficult detection case. This helped identify the strengths of each filter in a very demanding scenario. We defined small objects equivalent to cylinders with diameter of one inch or less. These objects simulated tree branches and other obstacles that have very little surface area. The sensors performed considerably better for medium and larger sized objects.

Figure 13 shows the response of both filters to a small object moving through the detection area. The Kalman filter provides the best combination of signal processing and lag time. However, the Gaussian filter is simpler to implement.

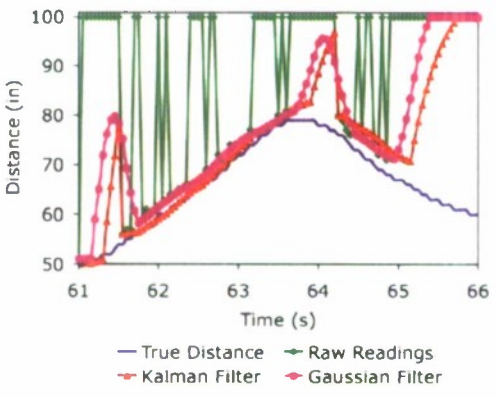


Figure 13 – Filter responses for a small object in the detection region.

Once sufficient data has been collected for both filters to operate, the filtered signals tolerate several missed readings before they deviate from the true distance. The Kalman filter follows the true distance measurement more accurately, and is more tolerant of missed readings than the Gaussian filter, reinforcing our conclusion that it is a better choice.

5. Filtering and Sensor Conclusions

Ultrasonic sensors provide a simple means of distance measurement ideally suited for use in mobile computing applications. When working with these sensors, it is important to recognize that their accuracy is dependent on the shape, size, and other physical properties of the objects to be detected.

Despite their simplicity, ultrasonic sensors can be difficult to work with because they are error prone in complex environments where object properties cannot be accurately predicted, or in cases where object location changes rapidly.

We were able to detect a wide variety of objects in our system by incorporating triangulation to better locate potential obstacles, and by filtering the results through a combination sliding window and Kalman filter. This combination allowed our system to tolerate unstable readings from small objects, as well as fluctuations in readings of objects near the periphery of the beam.

When the application is sufficiently simple, or the processing power available is especially limited (e.g., only limited mathematical operations available), a low-pass filter can be used instead. However, the use of a low-pass filter introduces a time delay into system response and decreases the accuracy of the result.

We used triangulation to increase the information available from sensor readings. Our method increases the available information, but it also complicates matters by reducing the polling frequency of each component sensor. We decided that, at this stage, the loss in polling frequency would not impede the design and testing of the system. The original polling frequency can be maintained through the use of frequency modulation for each sensor [3], or through other methods.

Triangulation can be used most effectively when the position of the sensor is known for each reading. In our case, since sensor location is arbitrary, we cannot accurately track objects within the detection field. This is because we cannot distinguish a moving object from a change in sensor position or rotation. Nevertheless, we could not track objects with a single ultrasonic sensor, so the additional information still enhances the capability of the system.

Our simulation was useful in the selection of sensors and the design of the system. These methods allow us to optimize distance and angle information given irregular coverage area data at the outer limits of the sensors. Simulations are not a valid replacement for actual test methods, but it gave us another way to measure our system's performance.

6. Tactile Feedback Implementation

While accurately tracking potential objects is important, without relaying this information in an intuitive manner, the filtering is irrelevant. The collision avoidance system for people with a visual impairment is separated into two main sections, filtering and feedback. With the filtered detection of the environment completed, the remaining task for the system is to intuitively relay the environmental information to the user. This section will focus on the practical implementation details and preliminary findings with using the tactile belt supplied by EAI, and will make suggestions for future research and uses. The full system is depicted in Figure 13; the batteries and control board are stored in a backpack.

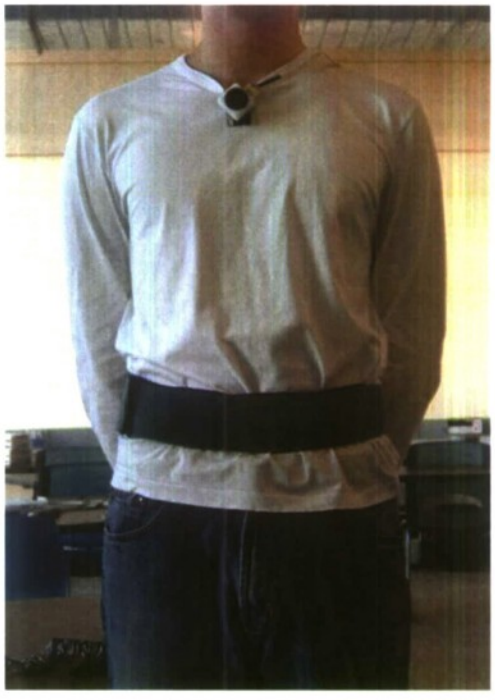


Figure 14 – Tactile belt and ultrasonic sensor (mounted at the neck).

Preliminary tests were performed on three subjects who were familiar with the project. The majority of tests were performed on a male subject who was tall (~6') and had an athletic frame (33" waist). The other two test subjects were used to test final encoding patterns as well as the difference between two different gain levels on the system. The first of these two remaining subjects was a female who was tall (~5'9") and had a medium build (36" waist). The second of the subjects was of medium height (~5'7") and had a medium build with a thin waist (30"). Since the belt is elastic in nature, it fit comfortably on each of the three subjects.

7. Preliminary Tactile Test Results

7.1 Tactile Frequency

To properly use the tactile information channel, the specific nature of the medium (the skin) must be better understood. Within human skin there are four main types of mechanoreceptors; the Meisner and Pacinan corpuscles, the Merkel disks, and the Ruffini endings [10]. Of these four mechanoreceptors, the Pacinan corpuscles (PC) are the most sensitive to vibrotactile stimulation. Just as there are visual and audio frequency ranges that humans can detect, there is a specific vibrotactile frequency range that the PC can detect. The PC mechanoreceptors detect frequencies in a range of 40-800 Hz in a U-shaped sensitivity pattern with a maximum sensitivity of 250 Hz. [10]

To ensure the best tactile response for the collision avoidance system, the tactile belt used within the system was tuned to 250 Hz. Preliminary tests used frequencies of 200 Hz, 250 Hz, and 300 Hz. No discernable difference was detected between these different frequencies, insinuating that there is a range of frequencies surrounding 250 Hz in which the tactile feedback will work and that the frequency does not need to be more finely tuned.

7.2 Gain Levels

In addition to the frequencies used, the gain (or strength) of the tactile response affects the performance of the feedback system. Maintaining the proper gain level is crucial to user satisfaction. The gain must be strong enough that the user will detect the tactile feedback while the user is walking. However, if the gain level is too strong the tactile sensation may be startling or unpleasant for the users as well as potentially damaging to the feedback system itself. Since the gain is directly related to the size of the amplitude of the signal sent to the tactors, an excessive gain for extended periods of time may either short or burnout the tactors by having too much current. The EAI belt comes with four gain setting levels, ranging from levels one to four, with gain level one being the weakest. Preliminary testing found that levels one and two were not detectable while wearing the belt at waist level, even if stationary.

Detection of the tactile signals began at gain level three; however, the signals were not consistently detected while the users were in motion. On this setting, detection also varied significantly depending on the amount of body fat where the tactors were located, as well as the placement of the tactors on the body. Of all the specific locations of tactors, users indicated that tactors resting above the pelvic bones were the most noticeable. Further investigation into tactor placement using the particular tactors employed by this system can help determine optimal tactor placement around the body in a pattern besides the current symmetric arrangement.

Tactors activated at gain level four were always detected during the preliminary tests; however, they caused users discomfort over extended use. Users in the preliminary tests indicated that when the front facing tactors (located above the abdomen) were activated for 3 bursts of 200ms with a frequency of 5 bursts a minute, the users experienced mild discomfort within 3 minutes of testing. These initial results indicate that gain level four, while highly detectable, should be used minimally to avoid discomfort for users. This gain setting may be further investigated to determine other factors to help minimize the impact of the discomfort, such as different pulse periods or different actuation frequencies.

7.3 Tactile Sequence Patterns

The collision avoidance system uses two different tactile sequences based on distance thresholds to potential objects. The first tactile sequence is activated at a distance of 50 inches. This tactile sequence is used as a last resort to warn the user of the impending threat, and thus is the strongest warning. This warning uses a gain level of four to make sure the user detects it. Additionally, this sequence activates the three most forward facing tactors to ensure that the user will detect the sequence on at least one of the tactors if the tactors are not located in the most ideal positions. Lastly, the tactors are activated for three bursts of 200ms on, 200ms off. A pattern of three short bursts each lasting 50ms was also used; the two sequences were equally effective.

The second tactile sequence is activated at a distance threshold of 75 inches. Based on the

ultrasonic detection data, obstacle detection at 75 inches is stable enough that it is unlikely to have many false positives. Additionally, it is an additional two feet farther out than the initial tactile sequence, allowing the user to be better notified of impending threats, adding extra safety to the system. Since this warning was for a less impending threat, it was designed to be less intrusive to the user. A gain level of three was chosen, and only one tactor was activated with a single pulse of 200ms. This tactile sequence was chosen to be interpreted as less significant and less imminent, while still notifying the user of a potential threat ahead. Additionally, since this tactile sequence would be activated more often, it was designed to limit the amount of tactile stimulation, as to help ensure that the user is not desensitized to the tactile feedback.

When testing burst duration, users did not prefer one burst length over another. While in general one encoding pattern was not more intuitive than another, users were able to determine that the signal that was longer, stronger, and used more tactors was more important than the shorter, weaker signal. Encoding patterns seem to become more important as more complex information is attempted to be passed through the tactile channel. The current encoding methods allow users to detect the relative difference between two different sequences; however the users are not able to determine specific distances to objects. When the results from the triangulation study are implemented into the system, location information will also be supplied by stimulating different tacots. More information on encoding patterns and tactile localization can be found in [8, 9, 11].

8. Conclusions

The collision avoidance system we have described was developed as part of an electronic travel aid, but the technology has many potential applications. The Navy has already explored utilizing tactile displays for helicopter hover tasks, and the U.S. Army has performed studies using tactile belts for waypoint navigation. Further studies on the ultrasonic sensors and tactile systems are currently underway. Undergraduate students are designing wearable systems with two ultrasonic sensors and a single tactor that will be field testing in the Spring 2008 quarter. These studies will determine if the algorithms developed in the current project can be applied outside the laboratory environment as part of a functional ETA.

9. References

[1] O.Wijk, P. Jensfelt and H.I Christensen. "Triangulation Based Fusion of Ultrasonic Sensor Data." *IEEE International Conference on Robotics & Automation, 1998.*

[2] Lei Fang, Panos J. Antsaklis, Luis Montestruque, M. Brett McMickell, Michael Lemmon, Yashan Sun, Hui Fang, Ioannis Koutroulis, Martin Haenggi, Min Xie, and Xiaojuan Xie. "Design of a Wireless Assisted Pedestrian Dead Reckoning System - The NavMote Experience." *IEEE Transactions on Instrumentation and Measurement, 2005.*

[3] Peter Veelaert and Wim Bogaerts. "Ultrasonic Potential Field Sensor for Obstacle Avoidance." *IEEE Transactions on Robotcis and Automation, 1999.*

[4] F. van der Heijden, P.P.L. Regtien. "Wearable Navigation Assistance - A Tool for the Blind." *Measurement Science Review, 2005.*

[5] Pitas, I., *Digital Image Processing Algorithms and Applications*, John Wiley & Sons Inc, ISBN: 9780471377399, 2000.

[6] Maybeck, P., *Stochastic Models, Estimations, and Controls*, Academic Press Inc., 1979.

[7] H.Z. Tan, R. gray, J. J. Young, and R. Traylor, "A Haptic Back Display for Attentional and Directional cueing", *Haptics-e: The Electronic Journal of Haptics Research*, Vol. 3, No. 1, pp. 20, 2003.

[8] Van Erp, J., "Vibrotactile spatial acuity on the torso: effects of location and time parameters", *IEEE First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems.* 2005

[9] Van Erp, J., Van Veen, H., Jansen, C., Dobbins, T., "Waypoint Navigation with a Vibrotactile Waist Belt", *Transactions on Applied Perception*, Vol 2, Sec 2., 106-117, 2005.

[10] Verrillo, R., "Investigation of some parameters of the cutaneous threshold for vibration", *Journal of the Acoustic Society of America*, Vol 34, No 11, 1962.

[11] Van Erp, J., Tactile *displays for navigation and orientation: perception and behavior.* ISBN: 978-90-393-4531-3. 2007.

This work was sponsored by the Department of the Navy, Office of Naval Research, under Award #N00014-06-1-1111.

Frequency Shifting Component for Miniaturized Laser Doppler Velocimeters

Project Investigator:

John Sharpe
Dynamical Systems Imaging Laboratory
Physics
California Polytechnic State University
San Luis Obispo, CA

The purpose of this projeet was to investigate a novel method for frequency shifting in laser Doppler velocimeters. This work should lead to low-eost, low-power, robust, directionally sensitive miniature laser Doppler probes that can be deployed in a number of commereial and research areas. The main aeeomplishments of the project are:

- The proposed eoneept has been proved by using an amplitude grating moved with a piezoeleetrie transdueer. Several technologies have been identified to implement the eoneept as a commereially feasible deviee.
- After a seareh of the literature and patents we have started filing a patent applieation on the idea.
- We have published a paper deseribing the system (*Sharpe JP. A phase-stepped grating technique for frequency shifting in laser Doppler velocimetry. Opt. Laser Eng. 45, 1067-1070 (2007))*
- The eoneept was ineorporated into the system design for a eompact, long-range velocity sensor and, in collaboration with an industry partner, Measurement Seience Enterprise, Ine. was submitted as a Navy STTR proposal to develop the sensor. The proposal was deelined.
- We have applied to NSF for funding to develop a high-speed switehable grating using liquid erystal teehnology. This proposal is pending.

Our immediate future plans for the project are:

- Refine our eomputer model for the system to investigate the use of pure phase-modulating gratings and the effect of multiple partieles in the measurement volume.
- Continue to seek industrial eollaborations and ways to commercialize the idea.

In the rest of this report we briefly deseribe the experimental work done to validate the eoneept and further details may be found in the publieation eited above.

A schematic of the test system is shown in figure 1.

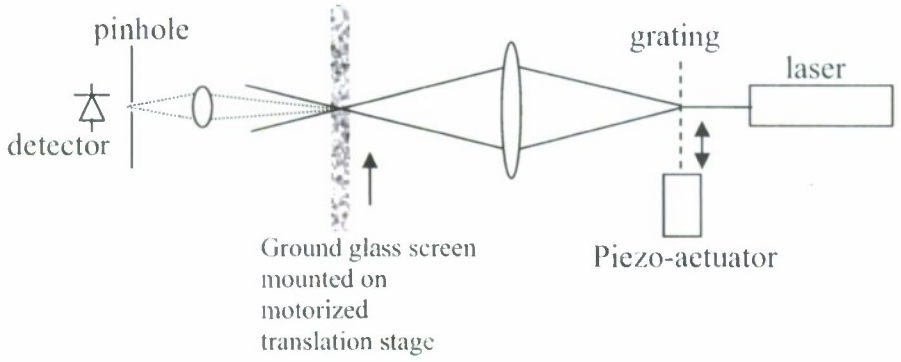


Figure 1: Schematic of arrangement to test the concept of discrete phase stepping in a laser Doppler system.

Light from a low power HeNe laser (λ=633nm) is incident on a 40lp/mm Ronchi grating (25 μm pitch). The +/- 1 orders are selected and the beams crossed on the ground glass screen with a grating magnification of ~ 1:1. Forward scattered light from the screen is collected and imaged through a 50 mm diameter pinhole to be detected. The optical fluctuation is recorded using a digital oscilloscope and transferred to a PC for processing.

The grating is mounted on a P.I. P733.2CL piezo stage and the stage driven with a P.I. E503 LVPZT amplifier The frequency response of this system is limited to several hundred hertz. We run the system in "open loop" to maintain speed.

Since the fringes are formed using the +/- 1 orders of the grating, a shift of half the period of the grating is needed to translate the fringes by one period – thus 12.5 μm. The algorithm calls for translations of 0, 1/3 and 2/3 of the fringe period, so the positions of the diffraction grating should be 0 μm, 4.2 μm and 8.3 μm. We programmed a function generator (Agilent 33120A) to supply the correct voltages to the PZT amplifier on which the diffraction grating was mounted. When driven slowly the required positions could be achieved to +/- 0.1 μm, though at faster speeds there appeared to be considerable amounts of ringing and overshoot. Thus we had to drive the stage at fairly slow speeds and we settled for 20Hz.

The data are shown in figure 1. (a) shows the spectrum when the ground glass screen is driven forward yielding a spectral peak at 5 Hz. This is reasonable since the stage moves at ~ 100 μm/s and the fringe spacing is 12.5 μm, so the nominal frequency is 8 Hz. In this case there is no movement of the grating. In (b) the screen is driven backwards and the peak frequency occurs at 7 Hz. The difference is peak position is because the translation stage moves slightly faster in the reverse direction than in the forward direction. In (c) the ground glass screen is stationary and the grating is oscillated at 20 Hz so there is a distinct peak at 20 Hz. In (d) the ground glass screen is driven forward so and the grating is oscillated. The predominant peak is at 25 Hz and stands about 10 dB above the next highest peak. (e) shows the case when the screen is reversed and the grating is oscillated and now the peak is at the expected value of 13 Hz.

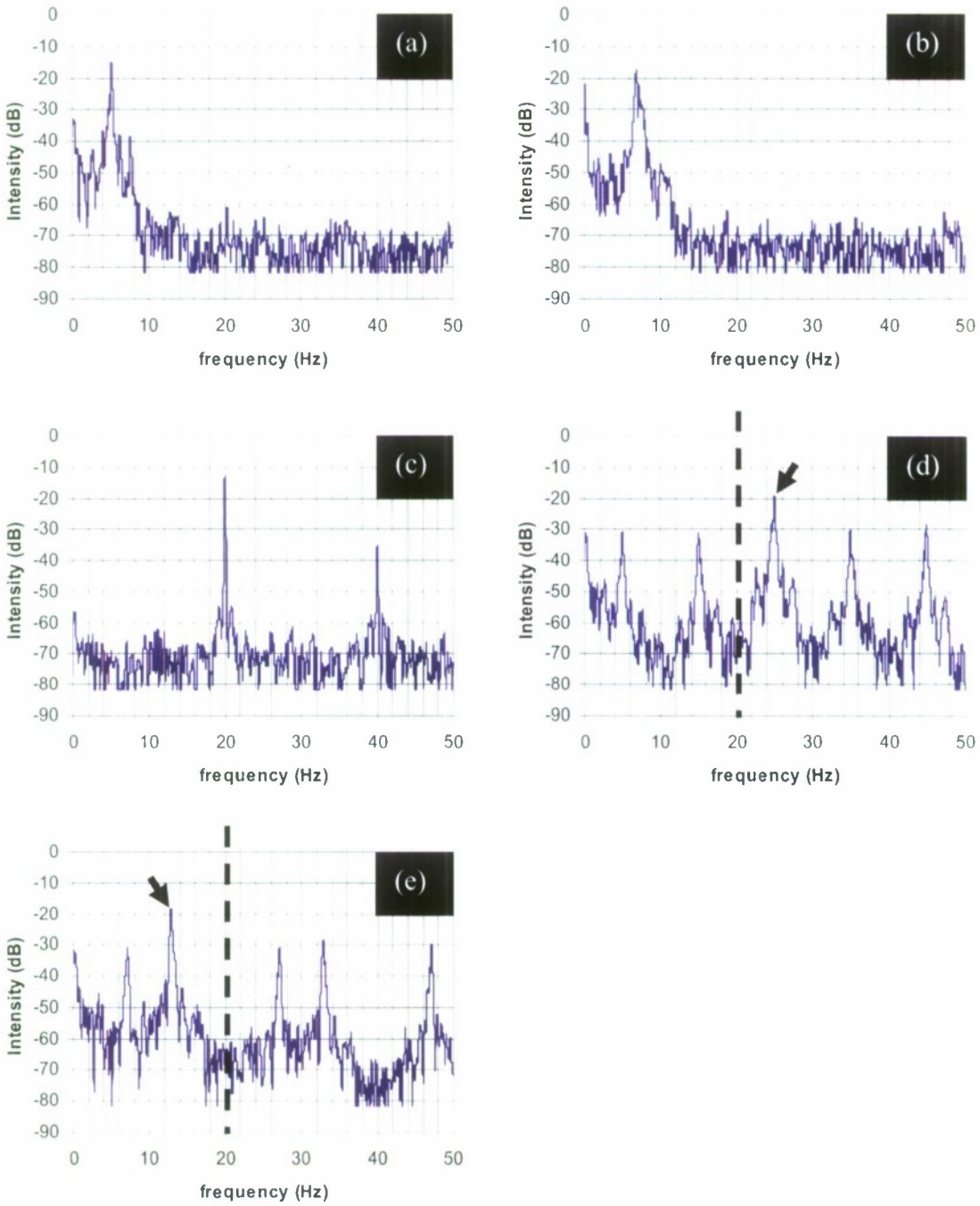


Fig 2 (a) Ground glass screen driven forward, peak at 5Hz. (b) Ground glass screen reversed, peak at 7Hz. (c) Screen stationary, grating oscillated at 20Hz. (d) Screen moved forward, and grating oscillated, peak at 25Hz. (e) Screen moved backward, and grating oscillated, peak at 13Hz

This work was sponsored by the Department of the Navy, Office of Naval Research, under Award # N00014-06-1-1111

A Computational Tool for Optical Forces

Project Investigator:

Nilgun Sungar
Physics
California Polytechnic State University
San Luis Obispo, CA

Project Objective

In this project our goal was to develop a computational toolbox that will allow the determination of the force on a particle trapped by a focused beam in the optical tweezers setup. This tool will allow computation of trap stiffness given the specification of particle structure and incident light field. Using this computational tool one will be able to study the effects of structure of the light field and particle shape and size that can aid in design of solutions for specific optical trapping problems.

Current state

During the spring and summer quarters of 2007, I worked two undergraduate students on this project, Angela Davidson and Dan Hernandez, both seniors majoring in physics. Dan graduated in August and is now employed by Heat, Light and Sound Research Inc. in San Diego, applying the same techniques we used on our project. Our results are summarized in the next section and with the renewal of our grant we will be able to expand the work as described in the last section of this report.

Results

The difficulty in the problem arose from two facts; the objective lens in the tweezers setup has high numerical aperture and the size of the particle in the light field is comparable to the wavelength of light used (~1μm). These limit the applicability of simplified approaches such as ray-optics or scalar Kirchoff beam propagation. To accurately model the propagation of light one needs to incorporate the full-vector nature of the electromagnetic field. We separated the problem into four distinct parts and formulated a strategy to optimize the computation of each part. These parts are

1- Representing the focused beam:

In traditional applications where the numerical aperture is small a Gaussian beam approach is used. Previous applications using this approach in optical tweezers produced results that are not in good agreement with the experimental measurements.
Our approach has been to combine the Ewald sphere /Pupil function approach [1] with the full vector diffraction theory [2]. In this approach the focused beam is represented as a sum of plane waves in the Fourier domain where the amplitude of each Fourier component is found using the pupil function. Each plane wave is then propagated using the vector diffraction method.

2- Propagating the beam through the cover slip and in the medium containing the trapped particles:

Passing of the focused beam through the glass cover slip produces spherical aberrations, changes the position of the focus and effects the polarization of individual plane wave components. Depending on the direction of the wave vector, the wave passes through a certain thickness of glass, which produces a phase-shift depending on the distance traversed in glass. Also one has to consider the transmission coefficients at each interface that can be found by the Fresnel formulas. In previous work by other groups [3], the effect of the cover slip was only accounted for by shifting the focal plane and changes to

the polarization of the field had been neglected. We have developed the theory for the passage of the plane wave components through the cover slip, including the polarization effect and implemented our results computationally. The codes we developed utilize the vectorized computation capabilities of the MATLAB programming platform for fast numerical integration. Numerical integration of the double integrals involved the use of quadrature method with implementation of the composite Simpson's rule. These computations allow for the adjustment of the numerical aperture of the microscope lens, cover slip thickness and index of refraction and the index of refraction of the trapping medium. Our results show that the electromagnetic fields in the trapping region have a rich structure as a result of the inclusion of polarization effects.

3- Finding the scattered electromagnetic field:

Once the field in the trapping region is computed, the insertion of a dielectric particle in the electromagnetic field gives rise to a scattered field. This scattered field produces a force (due to momentum change) that is in addition to the gradient force of the electric field. The total electric field around the trapped particle is then the superposition of the incident field and the scattered field. In order to calculate the scattered field we used Mie theory[4,5], where the electric fields and their derivatives for each plane wave component are matched at the boundary of a spherical particle. While Mie theory gives accurate results, it is limited to spherical particles. Although other approaches exist for arbitrary shaped particles we chose to use the Mie theory due to its accuracy and our current interest in extending the computation of forces in the presence of a secondary particle. We have written preliminary Matlab codes for the computation of the scattered fields based on Mie theory. These codes need refining and will require modifications to extend the current work to include a secondary particle. As explained in the future work section, during the spring quarter and summer of 2008 we will continue this project with funding through the renewal of our C3RP grant.

4- Calculation of the force

The most direct approach for determining the force vector is to use the total field at the location of the trapped particle and to calculate the Maxwell stress tensor [6]. However this approach is computationally very intensive and not practical for inclusion of a second particle in the vicinity of the trapped particle. Another approach is to calculate the gradient force and the scattering force separately [7]. It is this approach we take for computing the force. In this approach, the force on a dielectric particle is derived using the polarization of the particle P which is assumed to be proportional to the electric field in the medium. The force density f is then given as:

$$f(r,t) = [P(r,t)\nabla]E_m(r,t) + \frac{\partial P(r,t)}{\partial t} \times B_m(r,t)$$

where E_m and the B_m are the electric and the magnetic fields in the medium. This leads to two terms for the force density f, which can be interpreted as the gradient force and the scattering force components. The polarizability of the particle is found using the Clausius-Mossotti relation. Hence, calculating the force density is straightforward when the fields in the medium are known. We have developed some preliminary codes which use the fields in the vicinity of the trapped particle (computed as explained in the previous parts) to compute the force density which is then integrated over the volume of

the trapped particle. This gives us the net force vector on the trapped particle. These codes also need refining and modification to include a secondary particle.

While working on this project, some preliminary experiments done by Dr. John Sharpe in the Cal Poly Physics Department showed that the presence of a second particle (which is frequently used to tether the biological sample) interferes with the field of the trapping beam and changes the force on the trapped particle. This is a problem that has not yet been studied theoretically or computationally and its study would make an important contribution to the study optical tweezers. Inclusion of this effect requires the alteration of steps 3 and 4 outlined above. We have been awarded a renewal grant through C3RP program to study this effect both experimentally and theoretically. A brief description of this future work and how it ties into our existing results is explained in the next section.

Future work

During the spring quarter 2007, Dr. Sharpe and I will continue to work on the project with student involvement. The experimental work will be done by Dr. Sharpe at the existing facilities at Cal Poly. The theoretical part of the work requires modification of the procedure we outline in the previous section. The inclusion of a second particle comes into play in the third step of our work, with application of the Mie scattering theory. We will have to apply Mie scattering theory iteratively. First we calculate the scattered fields due to the secondary particle. For off-axis particles it is more practical to place the particle at the origin and shift the incident beam. This involves expressing the direction of each plane wave component of the incident beam in terms of Euler angles. Then, using structural information about the scatterer, boundary conditions are applied and scattering amplitudes are calculated for each scattering angle. After the calculation of the scattering amplitudes, the scattered field must then be shifted again to realign for the second iteration. At this stage the scattered field is combined with the incident field to obtain the net field and this field is scattered from the trapped particle. When the theoretical calculations are completed we will implement the results computationally. This implementation will require modification of the last two stages of the codes that are already developed. We expect to publish a paper at the completion of this work.

References cited

[1] A. Rohrbach and E. H. K. Stelzer, J. Opt. Soc. Am. A, 18 (4),839 (2001)
[2] S. Guha and G. Gillen, J. Opt. Soc. Am. B, 24 (1), 1 (2007)
[3] A. Rohrbach and E. H. K. Stelzer, Applied Optics, 41 (13), 2494(2002)
[4] H.C.van de Hulst, *Light Scattering from small particles*, Dover Publications, New York (1981)
[5] A. Rohrbach, Phys. Rev. Lett., 95, 168102 (2005)
[6] J.P. Barton and D.R. Alexander, J. Appl. Phys., 66(10), 4594 (1989)
[7] A. Rohrbach and E. H. K. Stelzer, J. of Appl. Phys. 91(8), 5474 (2002)

Establishment of an Electrophysiology Laboratory for Experimental Validation of a Combined Model of the Neuromuscular Junction under the Influence of Acetylcholinesterase Inhibiting Neural Toxins

Project Investigator:

Robert Szlavik
Biomedical and General Engineering
California Polytechnic State University
San Luis Obispo, CA

Project Objective

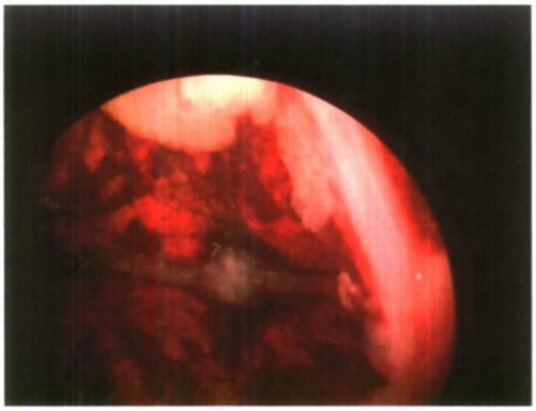
The major objeetive of this projeet was to establish a neurophysiology laboratory at California Polyteehnie State University. This facility would initially be used for implementation of a series of experiments to provide empirical data for validation of a theoretical model of the neuromuseular junction under the influence of aeetylcholinesterase inhibiting neural toxins. A theoretieal model has been developed by the principal investigator's researeh group that eneompasses aeetylcholine release from presynaptie vesicles, diffusion proeesses related to the neurotransmitter and organophosphate based neural toxins in the eleft and interaetion of the neurotransmitter with the postsynaptie nicotinie reeeptors at the motor end plate. The model provides a theoretieal predietion of the impaet of acetylcholinesterase neural toxins on the modulation of the postsynaptie musele fiber membrane conduetance at the motor end plate and concomitant modulation of the motor end plate currents.

Major Project Tasks

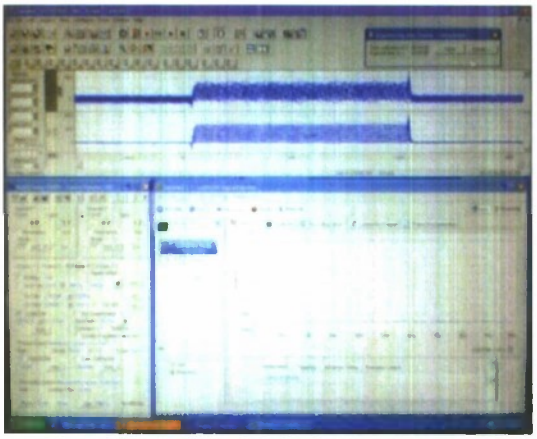
Task 1: Acquisition of Necessary Equipment and Supplies

To establish the electrophysiology/neurophysiology laboratory facility, the purchase of several major pieees of equipment and instrumentation were made. This equipment and these supplies have been purehased through funds from this grant as well as matching funds eommitted by the Department of Biomedieal & General Engineering. Some of the major equipment acquisitions are deseribed below.

- RF Sereen Room – An RF sereen room (Faraday Cage) facility has been installed. This facility provides a low eleetrieal noise environment for earrying out eleetrophysiology experiments.
- Eclipse FN1 Mieroseope – An eleetrophysiology microscope manufactured by Nikon has been installed and is eurrently in use.
- Electrophysiology Instrumentation Suite – A voltage/current/patch clamp amplifier digitizer from Moleeular Deviees (Axon Instruments MultiClamp 700B and Digidata 1440A) has been installed, calibrated, configured and is currently being utilized in related experiments.
- Micropipette Puller – A programmable mieropipette puller, P-97 from Sutter Instruments has been receeived, tested and is being utilized to pull sub mieron tip mieropipettes that are used in the experimental eomponent of this study.
- Miscellaneous Equipment and Supplies – Various miscellaneous items ineluding a mechanically stable mieroseope table, instrumentation raeks and a dissection microseope have been installed and are eurrently in use.



Leech Dissection Experiment Exposing One of the Ganglia Along the Nerve Trunk



Electrophysiology Instrumentation Suite Calibration Experiment

Task 2: Laboratory Setup and Configuration

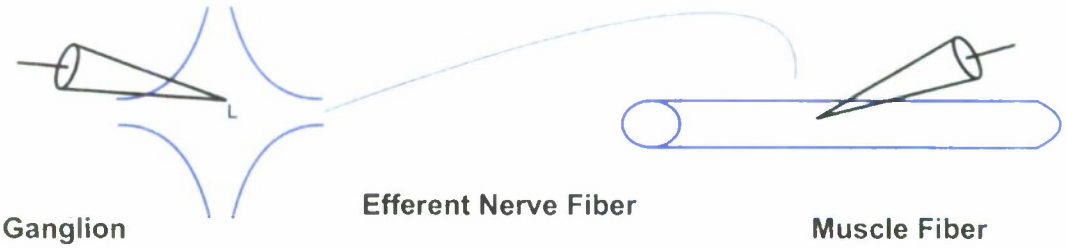
The various tasks related to the establishment of the laboratory facility have been completed.

- A graduate student (James Caulfield) was hired for assisting with the setup of the laboratory facility and assisting with the related experimental protocols. Mr. Caulfield has a background in Electrical Engineering and is in the Biomedical Engineering Master's Program. He has completed a course at Cal Poly on Biological Neural Simulation and Modeling (ENGR 550) that was given by the principal investigator during the spring quarter of 2007. Mr. Caulfield's Master's thesis dissertation will encompass the experimental studies that he will be performing related to this project.
- The student assistant has been trained in the use of the Molecular Devices electrophysiology instrumentation that has been acquired for this project.

Task 3: Experimental Study

The experimental component of this project has commenced. Preparations for the experimental work related to training and equipment configuration planning have been completed.

- The dissection protocols related to the leech animal model being utilized in the experimental studies have been successfully developed. These protocols are currently being utilized to prepare specimens for the electrophysiology studies that are currently underway.
- The first experiment will involve measurement of the time course of the motor end plate currents under a controlled condition (i.e. without exposure to the neural toxin) and after exposure to the neural toxin.



First experimental protocol will involve measurement of the time course of the motor end plate currents. A current clamp micropipette will be used to stimulate the L cell in the leech ganglion. The resultant motor end plate currents will be recorded at the postsynaptic muscle fiber under control conditions and after exposure to the neural toxin paraoxon.

- A second experimental protocol will involve measurement of the transmembrane potential at two points along the muscle fiber that are separated by a known distance. The transmembrane conductance per unit length will be computed from the solution to the time independent cable equation for intracellular excitation

$$v_m(z) = \frac{I_e \sqrt{r_m r_i}}{2} e^{-z\sqrt{\frac{r_m}{r_i}}}$$

Where I_e (A) is the intracellular injected current, z (m) is the distance between the intracellular current clamp electrode and the recording electrode, v_m (V) is the transmembrane potential and r_m and r_i are the membrane and intracellular resistance per unit length respectively (Ω/m).

Overall Project Goals in Relation to Naval Funding Objectives

This project has resulted in the establishment of an electrophysiology laboratory facility that is being used to undertake experimental studies on the impact of organophosphate based neural toxins on the neuromuscular junction. These studies are being used to empirically validate a theoretical model that the principal investigator's research group has developed that quantitatively describes the impact of organophosphate based neural toxins on the electrophysiology of the neuromuscular junction.

If this model can be successfully validated, then the potential exists to develop it into an equivalent circuit form whereby it could be utilized as a sub system in a neural toxin detection instrument for the purpose of determining the physiological impact of a given quantity of neural toxin on an organism. The model has further potential for utilization as a tool to determine the efficacy of therapeutic interventions for combating the effects of organophosphate based neural toxin exposure.

A further year of support has been secured that will fund these experiments to their conclusion.

Environmental Proteomics: The Minimal Stress Proteome in the Marine Model Organisms *Ciona intestinalis* and *C. savigni*-Networks of Co-Expressed Proteins

Project Investigator:

Lars Tomanek
Biological Sciences
California Polytechnic State University
San Luis Obispo, CA

RESULTS

During the last nine months we described the proteomic stress response of the marine sea squirt species *Ciona intestinalis* and *C. savignyi* in response to salinity and temperature stress. The objective of this project is to describe the **minimal stress response (MSR)** that is induced in response to the macromolecular damage caused by the acute exposure to multiple stresses. The distinction between the biochemical pathways common to all stresses as well as the ones that are specific to a particular stress provides valuable insights about the common and specific biomolecular targets of each stress. In addition, the characterization of the MSR leads to a better description of the behavior of system properties and their dynamics in response to perturbations through the environment. Such descriptions provide us with insights about the properties that determine the robustness and resilience of biological, specifically cellular systems.

Salinity stress

The characterization of the MSR requires a detailed analysis of the proteomic response of *Ciona* towards several stresses. We have chosen three of the most ubiquitous stresses: salinity, temperature and oxidative stress. In this first phase of this project we described the response to salinity (hyposaline conditions) and temperature stress.

The two *Ciona* species are found at pilings of harbors and marinas world-wide (*C. intestinalis*) and in the northeast Pacific (*C. savignyi*) where they are exposed to changing salinity conditions during heavy winter rains and subsequent run-offs (Dybern, 1965; Hoshino and Nishikawa, 1985; Lambert and Lambert, 1998). For example, in San Francisco Bay winter run-offs can lower the salinity of the water from 32 ‰ to 10 ‰. Massive die-offs of *Ciona* populations occur during these episodes of hyposaline conditions that are typical for the winter and they are followed by re-colonization in the spring (Lambert and Lambert, 1998). Personal observations in the field and embryonic studies (Dybern, 1967; Marin et al., 1987) suggest that the two *Ciona* species differ in their tolerance towards hyposaline stress and that it may be a contributing factor enabling *C. savignyi* to expand its distribution range faster than *C. intestinalis* (Lambert and Lambert, 2003).

Salinity changes cause massive cellular adjustments due to the osmotically driven influx of water during hyposaline conditions. The increase in cell volume that follows exposure to hyposaline conditions requires that the cellular scaffold be restructured (Pedersen et al., 2001). Active volume decrease is facilitated by the efflux of inorganic ions such as K^+ and Cl^- (Okada et al., 2001).

In order to characterize the changes in the proteome of *Ciona* in response to hypolsaline conditions we exposed both species to decreasing salinities (100%, 85% and 70%) for 6 h. Following the exposure to hyposaline (or –osmotic) conditions specimen were brought back to 100% salinity to recover for 4 h. A union fusion or proteome map of all the proteins detected through two-dimensional gel electrophoresis is shown for *C. intestinalis* in Fig.1.

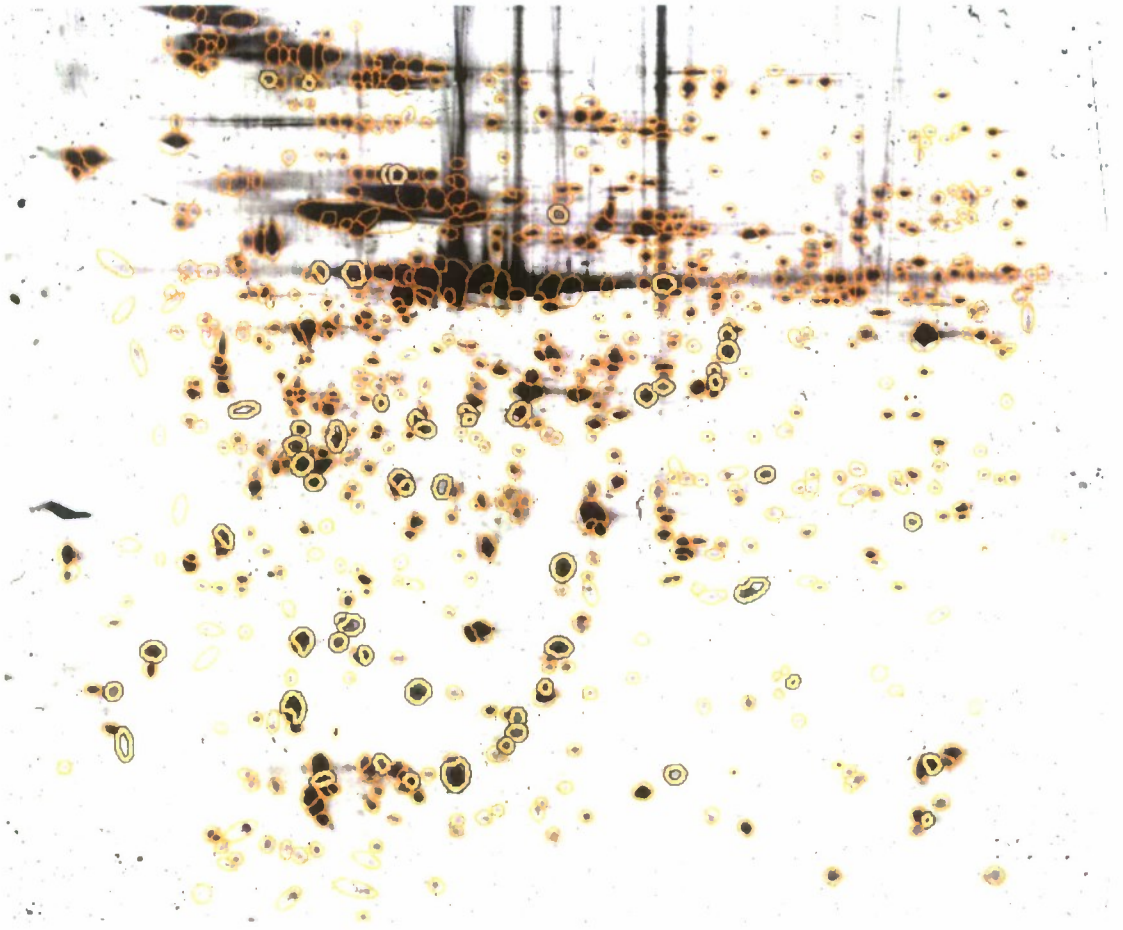


Figure 1: 'Fusion' gel that is equivalent to a proteome map of fifteen gel images (IPG gradient: pH4-7; length of IPG gel strip: 24 cm; Protein laoded: 800 µg) from whole animals (tunic removed) exposed to 100% (control), 85% and 70% salinity for 6 h followed by 4 h of recovery at 100% salinity (constant 13°C). We detected 727 protein spots. The group of spots selected in bold changed expression levels in response to either 85% or 70% salinity ($p < 0.05$). N= 5 for each salinity treatment.

Out of 727 protein spots 57 (7.8%) were differentially expressed in response to hyposaline conditions (t-test; $p<0.5$). Those that changed (increased or decreased their levels) are marked bold in Fig. 1.

As shown in the expression profiles in Fig. 2, levels of expression of 32 out of 43 proteins increased in response to exposure to 85% salinity but then decreased their levels in response to an exposure to 70% salinity. This general trend suggests that *C. intestinalis* is able to activate a cellular response to an acute stress to 85% but that the cellular machinery that supports this response is not able to activate the same defense response towards stress when experiencing greater salinity stress (70%). Although we have included *C. savigni* in our original salinity stress experiment, we have not finished the analysis of its protein expression profiles.

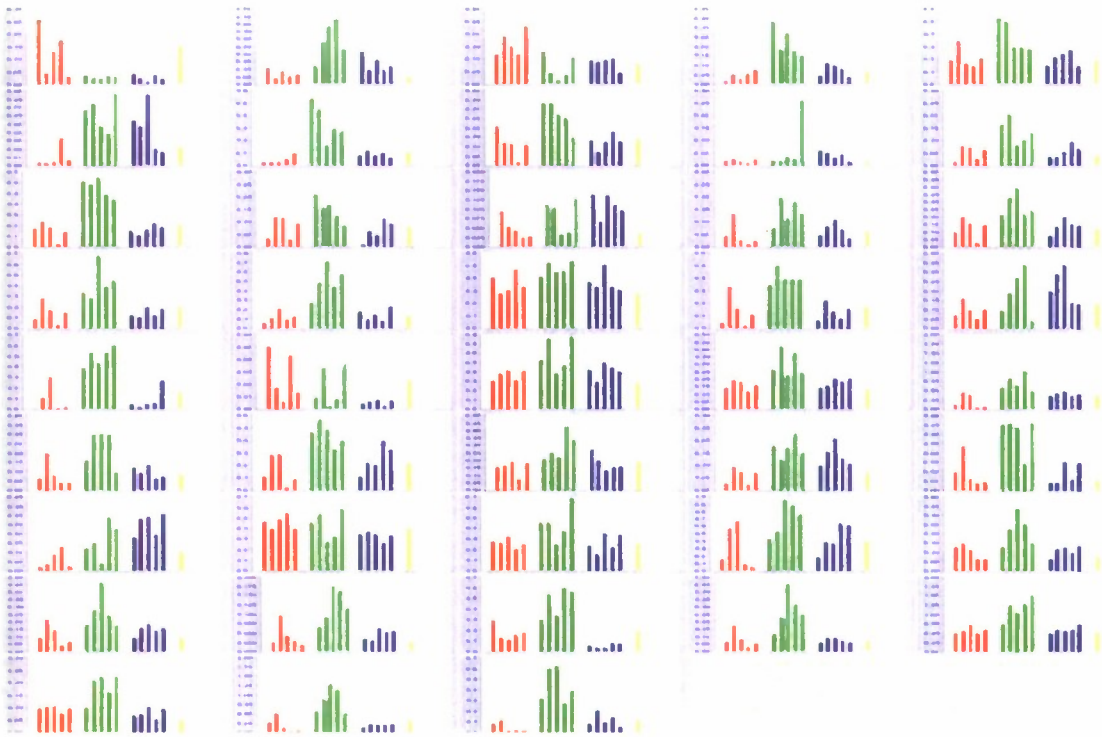
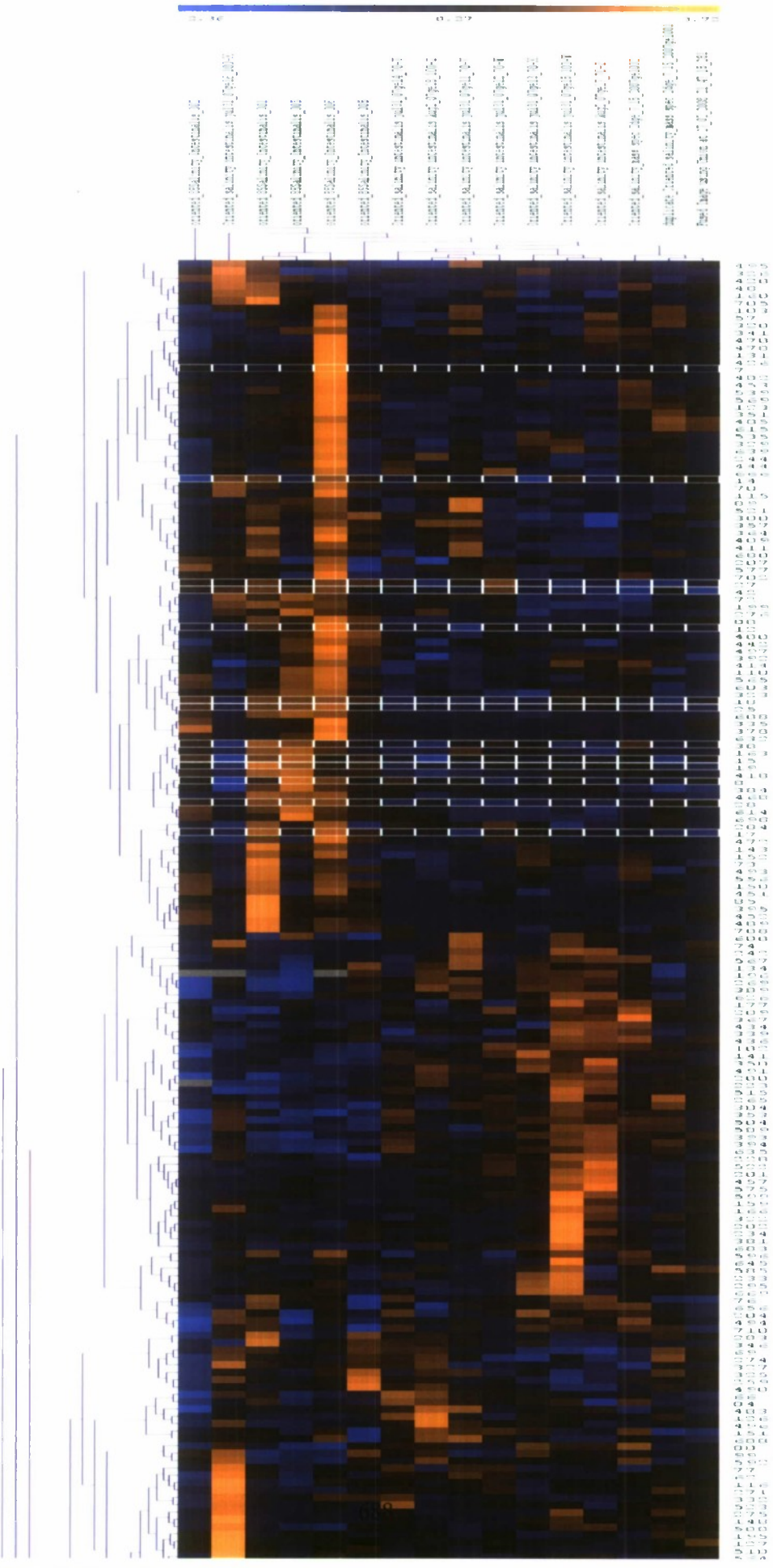
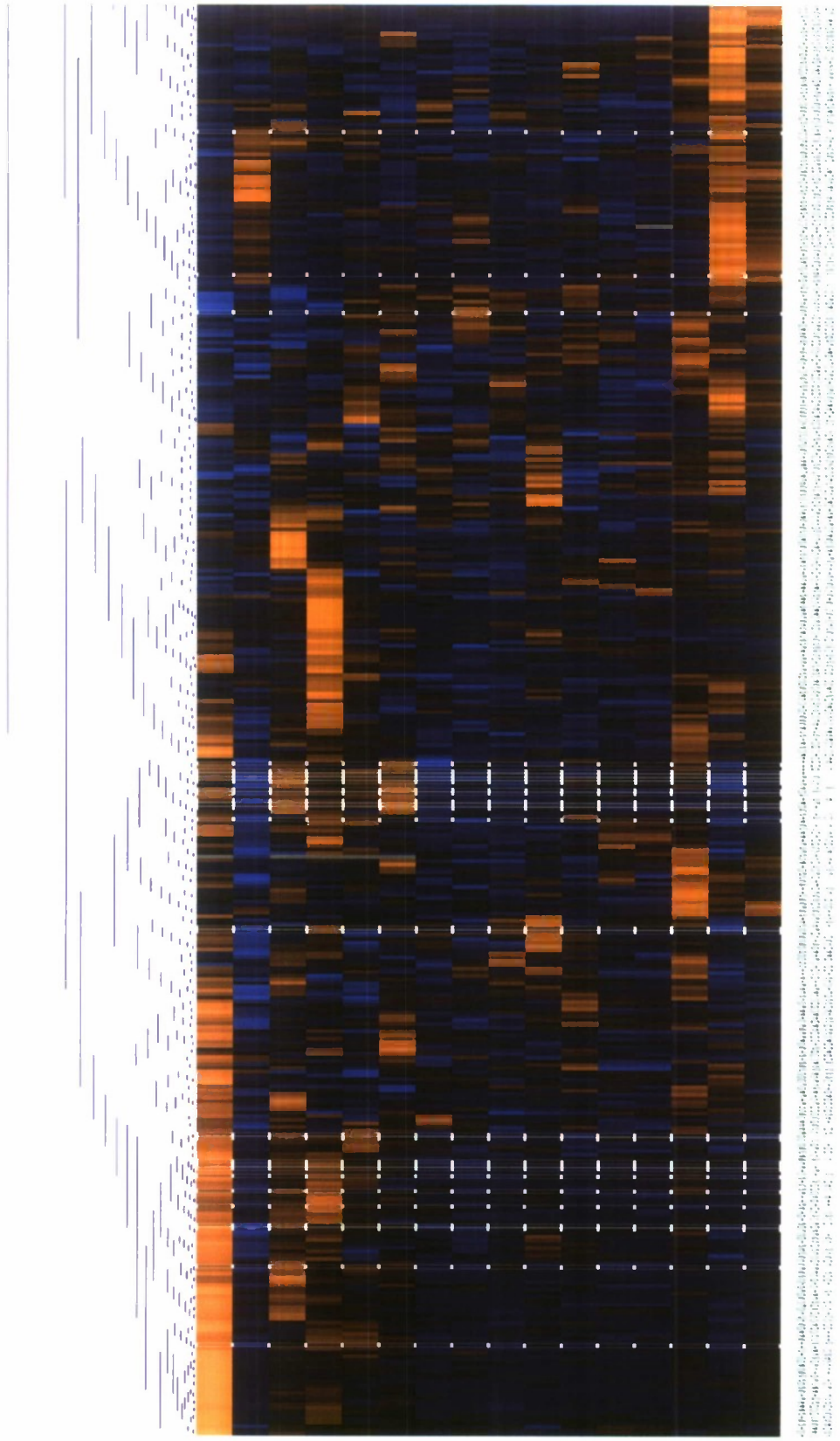


Figure 2: Expression profiles of 43 out of the 57 protein spots. Red = 100% salinity, green = 85% salinity, blue = 70% salinity and yellow = overall average.

In Fig. 3 we present a hierarchical cluster analysis of all the protein spots, their expression profiles with changing salinities and show how spots that differ significantly in their expression are found around clusters (shown in white lines surrounding the 'heat' or intensity squares). Proteins clustering close to each other are thought to be part of a common biochemical pathway. The analysis shows at least three major clusters that share between seven to ten protein spots. We are currently in the process of analyzing these spots and assign protein identities with the help of mass spectrometry. For a detailed description of the proteomic workflow we are applying see pages 3-4 in the original proposal.





A

Figure 3: Hierarchical cluster analysis of the salinity stress experiment for *Ciona intestinalis*.

Temperature stress

Temperature is arguably the most ubiquitous environmental factor influencing organismal biology, specifically in ectotherms that rely on ambient temperatures (Hochachka and Somero, 2002). For example, it is one of the main factors determining the distribution range of marine invertebrates, such as intertidal molluscs (Tomanek, 2002; Tomanek, 2005; Tomanek and Sanford, 2003; Tomanek and Somero, 1999). Adjustments in global gene expression levels are widespread with varying temperatures (Podrabsky and Somero, 2004). The two *Ciona* species differ in temperature tolerance, although it is still unclear to what degree. We proposed to elucidate their proteomic response to acute and chronic exposures to a wide range of water temperatures. Here we present the proteome map of all the protein spots that we detected in *Ciona savignyi* (1549) and in bold show protein spots that had statistically different (t-test; $p < 0.05$) expression levels following an acute exposure to 30°C for 2 h with 6 h of recovery at 13°C (Fig. 4).

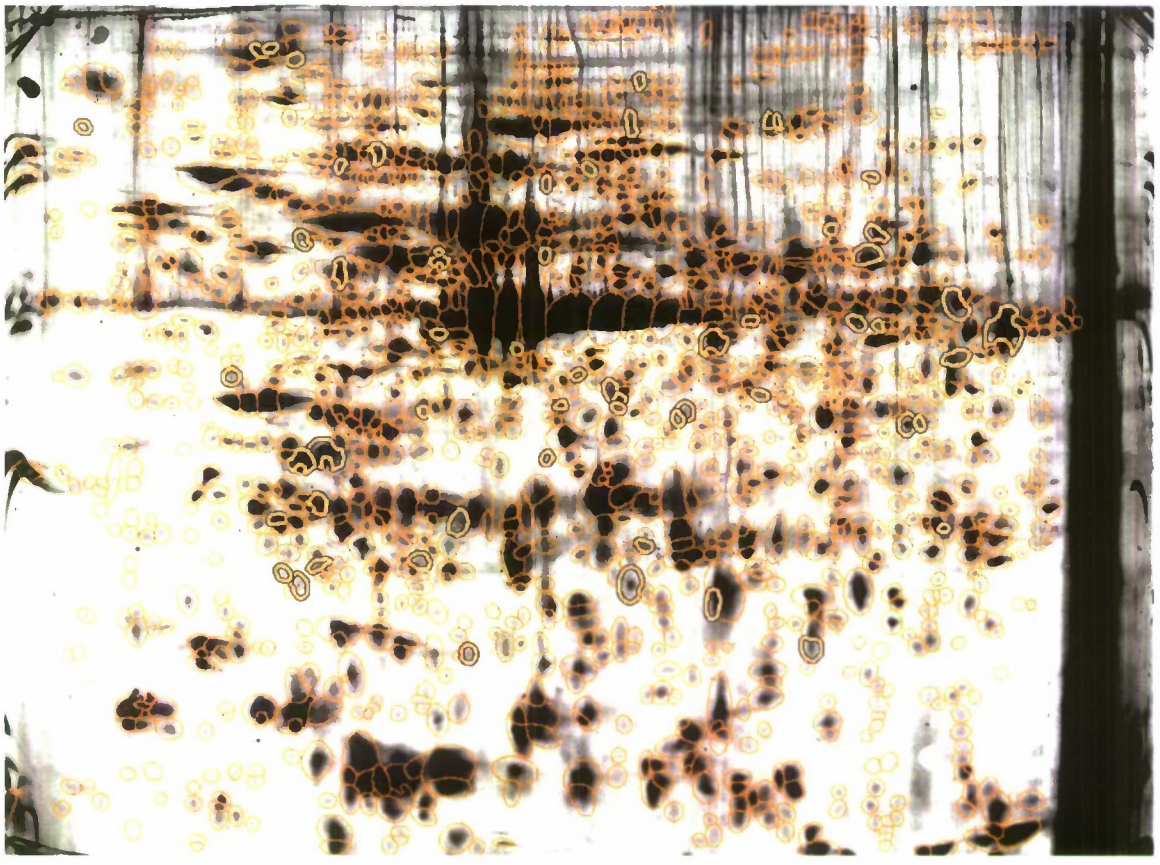
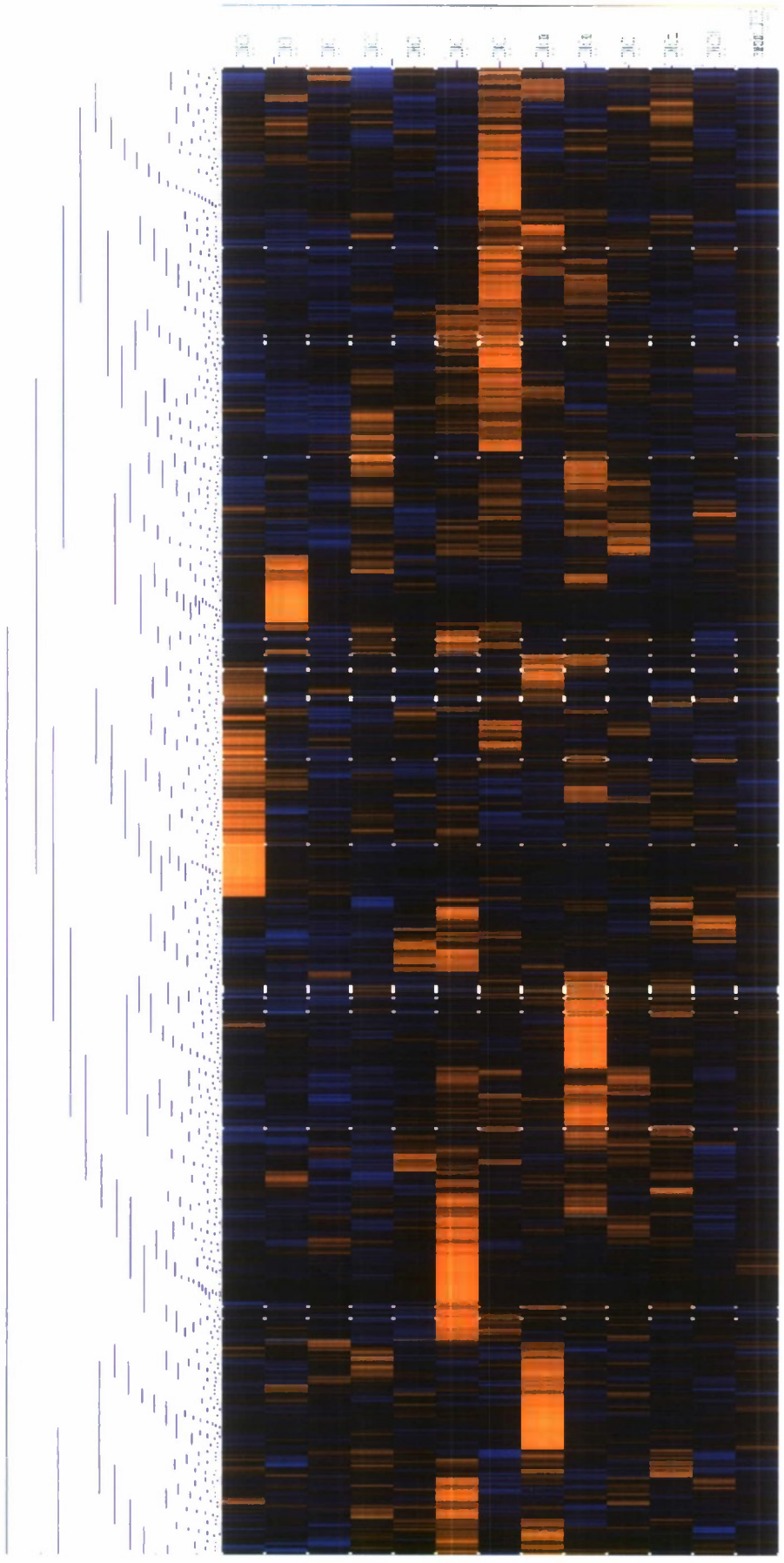


Figure 4: 'Fusion' gel of twelve gel images (IPG gradient: pH4-7; length of IPG gel strip: 24 cm; Protein loaded: 800 µg) from whole animals (tunic removed) exposed to 13°C (control), 22°C, 26°C and 30°C for 2 h followed by 6 h of recovery at 13°C. We detected over a thousand protein spots. The group of spots selected in bold changed expression levels in response to the acute exposure to 30°C in comparison to the 13°C control ($p < 0.05$). N= 3 for each heat treatment.

Figure 5: Expression profiles of 58 out of the 1549 protein spots that showed statistically different expression levels after incubation to 30°C in comparison to the 13°C control. Red = 13°C, green = 22°C yellow = 26°C, blue = 30°C and white = overall average of all gel images and temperatures.

Expression profiles of 58 protein spots that showed statistically different changes in expression are shown in Fig. 5. Two general patterns emerge: 15 protein spots increase their expression levels with heat exposure, suggesting that they play an important role in the cellular response towards thermal damage. Almost all the other proteins show a down-regulation of expression, a typical pattern for proteins that are not part of the cell's heat-shock response (Tomanek, 2005; Tomanek and Somero, 1999).



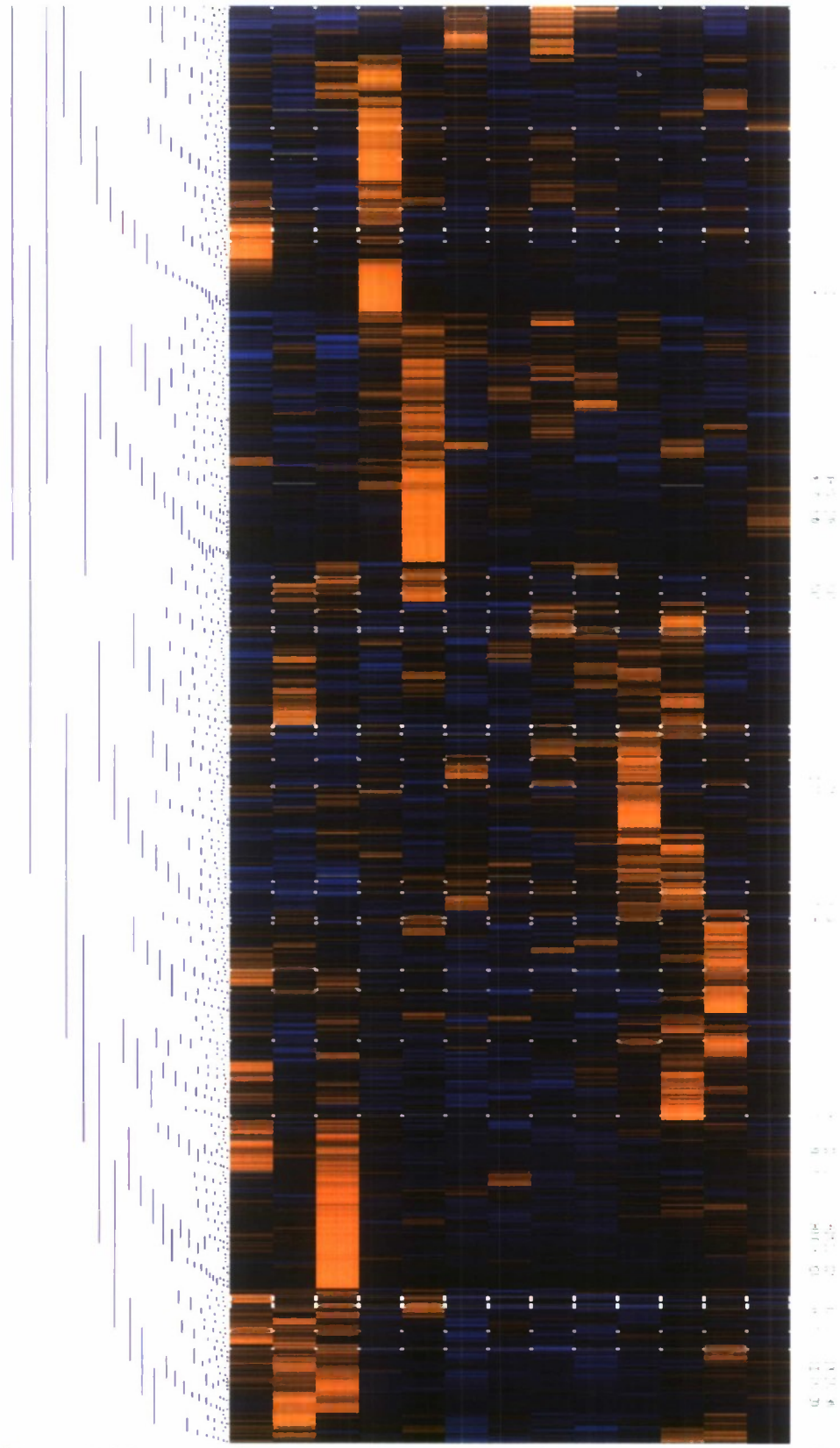


Figure 6: Hierarchical cluster analysis of the heat stress experiment for *Ciona savignyi*.

The cluster analysis of protein expression patterns shows that spots that change expression levels are distributed throughout the heat map of Fig. 6. This is most likely due to the fact that the down-regulation of non-heat shock proteins is non-selective. However, a cluster of proteins should exist that potentially represents heat-shock proteins. Fig. 7 shows a close-up of our analysis. The pattern shows six protein spots that are up-regulated at higher temperatures and that are clustering together, suggesting that their function is either similar or that these proteins share a common biochemical pathway that is activated in response to heat.

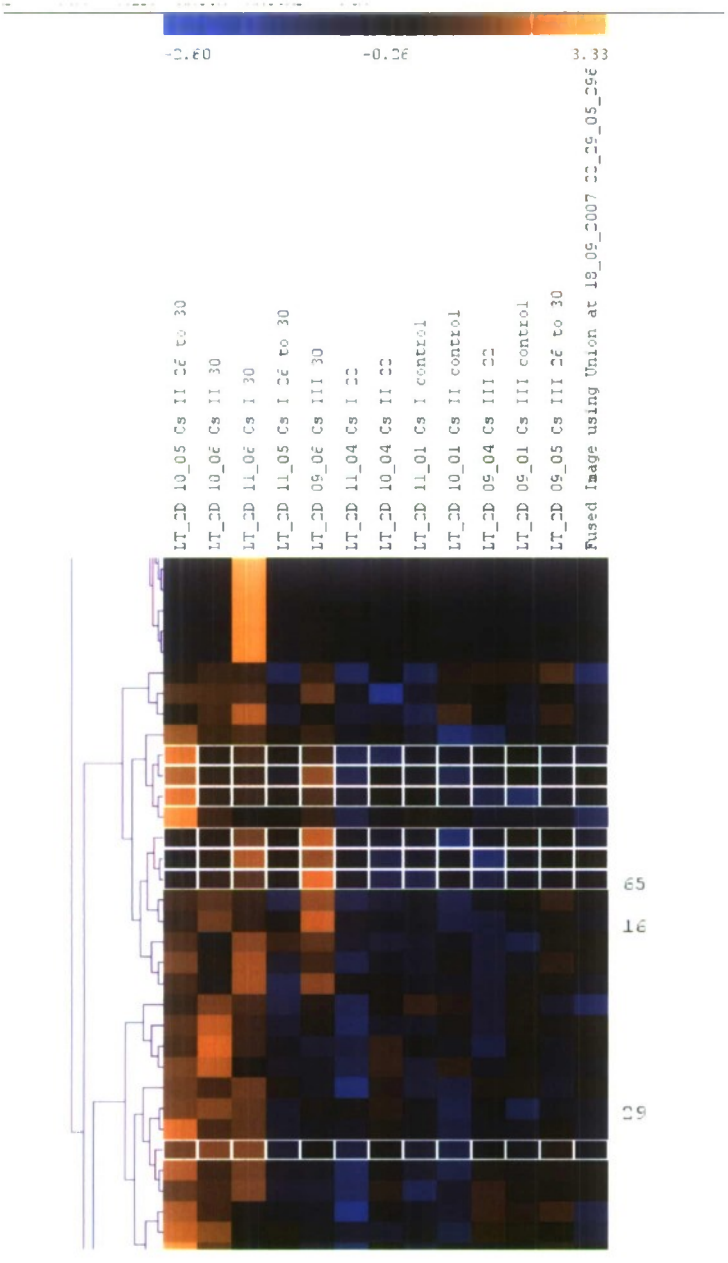


Figure 7: Close-up of cluster analysis showing an expression profile of heat-induced proteins.

Mass spectrometry and identification of proteins spots

Although we have been successful at conducting mass spectrometry (Fig. 8) and have identified several protein spots (Hann and Tomanek, 2007), we have not been able to complete the analysis of proteins that are of interest to us due to their changing expression patterns in response to stressful conditions. We are currently in the process of completing this part of the project.

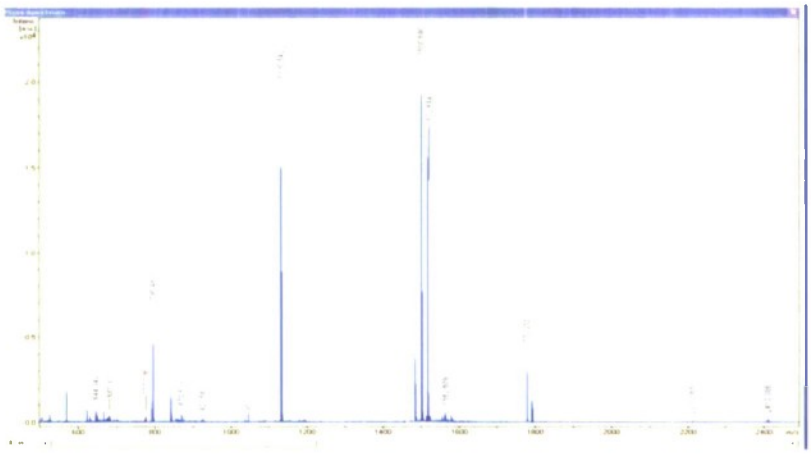


Figure 8: Distribution of peptide masses for *Ciona* spot 'C2'. This peptide mass fingerprint is unique and provides us with the identity of the protein (after Hann and Tomanek, 2007).

SUMMARY

Our results show for the first time how a marine invertebrate that is closely related to vertebrates responds to temperature and salinity stress at the proteomic level. The emerging patterns of protein expression present clusters that we hypothesize are proteins that are related by function or biochemical pathway. The methodology that we applied opens a way to a systems biology approach to study environmental stress in marine organisms. The resulting protein expression profiles may be used as indicator for specific environmental conditions. This approach can be used in the future to develop biomarkers or –sensors that are indicators for a range of environmental stresses, e.g. physical, chemical to biological. Our next major step is to identify the proteins through mass spectrometry, assign functions to the identified proteins, compare pathways that they share and develop new hypotheses of how cells respond at the systems level to environmental stress.

ACKNOWLEDGEMENT OF SUPPORT

This work was supported by the Department of the Navy, Office of Naval Research, under Award #N00014-06-1-1111.

REFERENCES CITED:

Dybern, B. I. (1965). The life cycle of *Ciona intestinalis* (L.) f. *typica* in relation to the environmental temperature. *Oikos* **16**, 109-131.

Dybern, B. I. (1967). The distribution and salinity tolerance of *Ciona intestinalis* (L.) F. *typica* with special reference to the waters around southern Scandinavia. *Ophelia* **4**, 207-226.

Hann, J. and Tomanek, L. (2007). Detecting osmotic stress proteins of *Ciona intestinalis* through functional proteomics. In *Annual meeting of the Western Society of Naturalists*, (ed. Ventura, California.

Hochachka, P. W. and Somero, G. N. (2002). Biochemical adaptation: Mechanism and process in physiological evolution. Oxford: Oxford University Press.

Hoshino, Z.-I. and Nishikawa, T. (1985). Taxonomic studies of *Ciona intestinalis* and its allies. *Publication of the Seto Marine Biology Laboratory* **30**, 61-79.

Lambert, C. C. and Lambert, G. (1998). Non-indigenous ascidians in southern California harbors and marinas. *Marine Biology* **130**, 675-688.

Lambert, C. C. and Lambert, G. (2003). Persistence and differential distribution of nonindigenous ascidians in harbors of the Southern California Bight. *Marine Ecology Progress Series* **259**, 145-161.

Marin, M. G., Bressan, M., Beghi, L. and Brunetti, R. (1987). Thermo-haline tolerance of *Ciona intestinalis* (L., 1767) at different developmental stages. *Cahiers de Biologie Marine* **28**, 47-57.

Okada, Y., Maeno, E., Shimizu, T., Dezaki, K., Wang, J. and Morishima, S. (2001). Receptor-mediated control of regulatory volume decrease (RVD) and apoptotic volume decrease (AVD). *Journal of Physiology* **532**, 3-16.

Pedersen, S. F., Hoffmann, E. K. and Mills, J. W. (2001). The cytoskeleton and cell volume regulation. *Comparative Biochemistry and Physiology. Part A, Molecular and Integrative Physiology* **130**, 385-99.

Podrabsky, J. E. and Somero, G. N. (2004). Changes in gene expression associated with acclimation to constant temperatures and fluctuating daily temperatures in an annual killifish Austrofundulus limnaeus. *Journal of Experimental Biology* **207**, 2237-54.

Tomanek, L. (2002). The heat-shock response: its variation, regulation and ecological importance in intertidal gastropods (genus *Tegula*). *Integrative and Comparative Biology* **42**, 797-807.

Tomanek, L. (2005). Two-dimensional gel analysis of the heat-shock response in marine snails (genus *Tegula*): interspecific variation in protein expression and acclimation ability. *Journal of Experimental Biology* **208**, 3133-43.

Tomanek, L. and Sanford, E. (2003). Heat-shock protein 70 (Hsp70) as a biochemical stress indicator: an experimental field test in two congeneric intertidal gastropods (genus: *Tegula*). *Biological Bulletin* **205**, 276-84.

Tomanek, L. and Somero, G. N. (1999). Evolutionary and acclimation-induced variation in the heat-shock responses of congeneric marine snails (genus *Tegula*) from different thermal habitats: Implications for limits of thermotolerance and biogeography. *Journal of Experimental Biology* **202**, 2925-2936.

PVD/CVD Coatings for Improved Life of Nano-Grain Cutting Tool for Machining Aerospace Alloys

Project Investigator:

Daniel Waldorf
Industrial Engineering
California Polytechnic State University
San Luis Obispo, CA

Summary

High-strength, heat-resistant metal alloys such as titanium and nickel-based superalloys continue to make up the biggest challenge in terms of economical production of metal components for the aerospace industry. These alloys show unusual work hardening and generate extreme temperatures during machining. As such, only expensive ceramic cutting tools typically last more than a couple of minutes before wearing out. Although cutting loads must be kept low, the ceramics can achieve high overall material removal rates due to the high speeds and temperatures at which they can maintain their edge. The first year of this C3RP grant focused on new tool material formulations for cutting superalloys with the goal of finding a lower cost tool that can match the life and material removal rates of ceramics. The study examined high-temperature, alternative materials (rhenium and superalloy) in addition to the standard cobalt as the binder constituents in a tungsten-carbide-based metal-matrix composite tool. The results from the past year show a clear advantage of the new compositions, with tools lasting up to twice as long as traditional carbides. The current research expands on those results by examining the effect of adding PVD and CVD coatings to further extend the lives of the new tools. The findings from additional cutting experiments on a new set of coated tool samples indicate a dramatic boost from the coatings. Tools show approximately 2-3 times the tool life as compared to the uncoated tools.

Plan of Work

A series of fabrication tasks and experiments was proposed to evaluate the effects of adding coatings to improve the tool life of the new alternative-binder cutting tools. As in the first year of funding, material samples were prepared by the project collaborator, Dr. Scott Liu of Genius Metal in Monrovia, CA. Of the tool formulations used in the first project year, the top five performing grades were selected to be coated for continued evaluation in this phase of the project. Although it was originally planned to coat one of the cobalt-only grades (i.e., d1, d2, or d3), the results of the first project year clearly demonstrated that the cobalt-only grades had inferior performance (approximately ½ the tool life) compared to the alternative-binder grades. The research therefore focused on the alternative binder grades and the specific effects of the various formulations within those grades. Table 1 shows the grades evaluated in the project. R-95 and U-720 are two different Ni-Based Superalloys. T3.5 and T32 are WC powder with the average grain size of 0.55 and 2 μm, respectively. As is typical with carbide tooling, small amounts of

	Re	R-95	Co	U-720	T3.5	T32	TiC	TaC	Cr_2C_3	VC	Density
	%	%	%	%	%	%	%	%	%	%	g/cc
d13	1.5	3	2.5	3	89	0			0.7	0.3	15.004
d15	0.75	1.5	1.25	1.5	94	0			0.7	0.3	15.27
d21	0.75	2.125		2.125	88	0	3	3	0.7	0.3	14.904
d22	1.5	3	2.5	3	83	0	3	3	0.7	0.3	14.648
d24	0.75	1.5	1.25	1.5	88	0	3	3	0.7	0.3	14.915

Table 1 Composition by Volume % of Tool Grades Used for Testing

chromium earbide (Cr_2C_3) and vanadium carbide (VC) were used in the formulation of the carbide phase.

The grades were selected by considering which inserts performed the best in the uncoated tests and by eonsidering a logical set of grades on whieh to base a eontrolled experiment. Sinee d13, d15, d22, and d24 all performed reasonably well in the uncoated tests, they were selected for eoating. They also form a well-defined 2x2 experimental design to test the effects of binder % (10% binder for d13 and d22, 5% binder for d15 and d24) and the presence of TiC and TaC in the carbide phase (d22 and d24 have TiC/TaC, d13 and d15 do not). Since d21 also performed consistently well in the uncoated tests and it provides a direct comparison with d24 with regards to binder composition, it was also chosen for coating.

Standard shape eutting inserts were fabricated from powder for each grade. The powder was purchased from an outside vendor (e.g., rhenium powder from Rhenium Alloys, Inc., in Ohio) and then subjeeted to the following proeessing sequence at Genius Metal:

- Milling (to reduee grain size), drying, and sifting the powders to ereate the right blend for each lot.
- Cold pressing in bulk using a hydraulie press under high pressure to form the initial "green" compact bars.
- Presintering in a conventional oven just hot enough to start the particle bonding that allows signifieant handling of the eompaets without risk of damage. 4 test bars for each lot were pressed and sintered in this manner.
- Additional cold pressing into the shape of cutting tool inserts (industry designation SNG-432), with around 20 inserts for each lot.
- Solid-phase sintering in a eonventional oven. This sintering does not reach the melting temperature of the metal matrix material (unlike with traditional sintering of cobalt-based carbides). The purpose is to further densify the samples and begin the real bonding between all of the constituent partieles.
- Hot isostatie pressing (HIP) to eomplete the bonding and densification. HIP involves placing the samples on a graphite plate and submerging in a high-temperature, high-pressure gas (typically Argon) medium. Here, the further physieal eompaetion and the high-temperature sintering happen simultaneously to aehieve the final material composition.

Samples were then sent to a machine shop for grinding and honing of the final surface, profile, and edge condition. Surface roughness is critical to fracture toughness in cemented carbides. Similarly, a smooth profile (e.g., corner rounding on square inserts), and appropriate edge condition ensure that premature fracture is avoided. For the hard, potentially brittle samples developed here, a small, rounded hone (approximately 0.01 mm radius) was put on each insert to provide the best defense against chipping and fracture.

Finally, the samples were sent to be coated. Ti-Coating in Utica, Michigan, was selected as the coating vendor. Two different coating types were selected, and roughly half of the tool samples produced were coated with each type. The first type was a CVD (chemical vapor deposition) coating grade denoted Al_2O_3, but really composed of a 9 μm thick two-layer combination of TiC and Al_2O_3. The second type is referred to as TiAlN (trade name Ti-Rex4) and is a 3um thick PVD (physical vapor deposition) coating of just TiAlN. Both of these coatings are commonly used to provide a barrier against the kind of chemical and thermally-induced wear experienced when cutting superalloys, but the TiAlN type has become the industry standard for cutting Inconel 718.

The samples were tested by running machining experiments in the Industrial and Manufacturing Engineering Department's Advanced Manufacturing Lab. Tests were carried out on a Haas HL-20 machining center, capable of 3400 RPM spindle speed. Cutting speeds of 50 and 100 meters per minute were used. Depths of cut and feeds were modest (2 mm depth, .1 mmpr feed). Tool geometry included negative rake angle inserts (SNG-433) in a MSRNR 16-4D 15° lead tool holder. No cutting fluid was used. Work material was as-forged, solution-treated Inconel 718 (150 mm diameter bar stock) purchased from Alloy Metals (Michigan). Measured tensile yield strength was 1,020 MPa and ultimate tensile strength was 1,120 MPa (hardness 36 HRC).

Figure 1 shows the machine set-up for the tests. The tests were randomized so that bar position and diameter would have a random effect. A fresh edge was used for each test. An automated vision system (OGP Smartscope 250) was used to examine wear modes and areas. It was used to take measurements of flank wear land width (VB) since flank wear dominated tool life for all of the conditions run. Three undergraduate student assistants (Mike Stender, Dan Norgan, Holly Dunekel) were employed to help with the experimentation by preparing work materials, programming the machine tools, collecting data, and assisting with technical problems that arose during the tests. Tests included short-time tests, with three replications made for each test condition, and run-to-failure tests on selected cutting edges. A commercially-available standard C2 grade, TiAlN-coated carbide insert designed for cutting superalloys (Valenite VC912) was employed for comparison in one of the run-to-failure tests.

Results

Results from coated tools with alternative binder build on the results from the first year of funding and are presented below. Those earlier tests on uncoated tools showed that tool grades with alternative binders (i.e., those that include rhenium or superalloy or both as

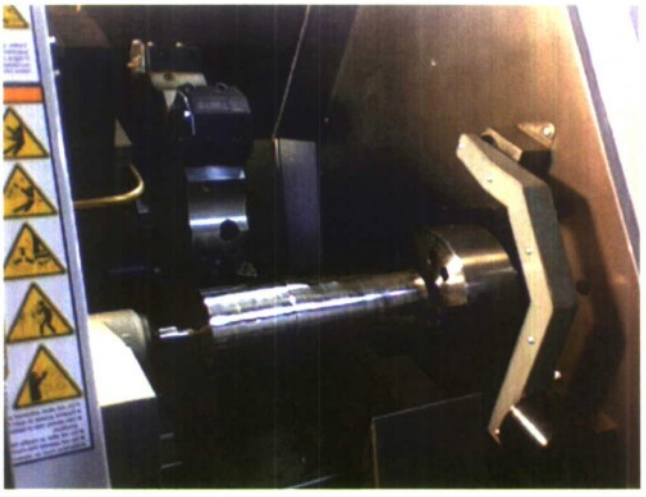


Figure 1 Lathe Set-Up for Tool Wear Test

binder metals) performed signifieantly better than tools bound only with cobalt, with the best tools lasting approximately twice as long. The earlier tests also showed a small beneficial effect (15% inerease in tool life) of adding TiC and TaC to the composition of the earbide phase in the tool and of restricting the overall % of binder in the tool to 5%. No clear results could be drawn from the uncoated tests about the optimal pereentage of Co, Re, or superalloy or about the effect of grain size.

Short time tests.

Figure 2 shows four graphs for the alternative binder tools, comparing uncoated results with tools coated with Al_2O_3 and TiAlN at two different times of cut, all at a eutting speed of V = 50 mpm. Results are from three replications for each test condition and represent comparisons at the same bar position and with the same edge eondition on each sample tool. In all cases, the eoated tools perform mueh better than the uncoated tools, exhibiting one-half or less of the wear level in virtually all eases, suggesting at least 2-3 times the tool life. The tests also show the TiAlN-coated tools performing eonsistently better than those coated with Al_2O_3 – with a wear level of about 50-75%. The tests, however, were all completed at relatively low levels of wear (less than VB = 0.1 mm in most cases), sinee a typieal eriterion for tool life is VB = 0.3 mm. Therefore, the conclusion that the TIAlN-coated tools perform better must be qualified by considering only the early stages of wear land development.

The same data is re-presented in Figure 3 with different vertieal seales to show the differenees among tool grades for each type of eoating. The tool grades include differenees in binder proportion (d13 and d22 have 20% binder, d14 and d24 have 5% binder) and in the presenee of TiC and TaC in the earbide phase (d22 and d24 have TiC/TaC, d13 and d14 do not). Both the uncoated tests and the TiAlN-coated tests show a significant effeet of the presence of TiC and TaC (at 5% significanee level). However, the effect only at the 70 seeond time for the uncoated tests and in both cases is a negative effeet; i.e., the presenee of TiC and TaC produees worse performanee, which is the opposite finding from the first round of experimentation. No significant effeet of TiC/TaC presenee is shown for the Al_2O_3 coating. None of the three eoating eonditions

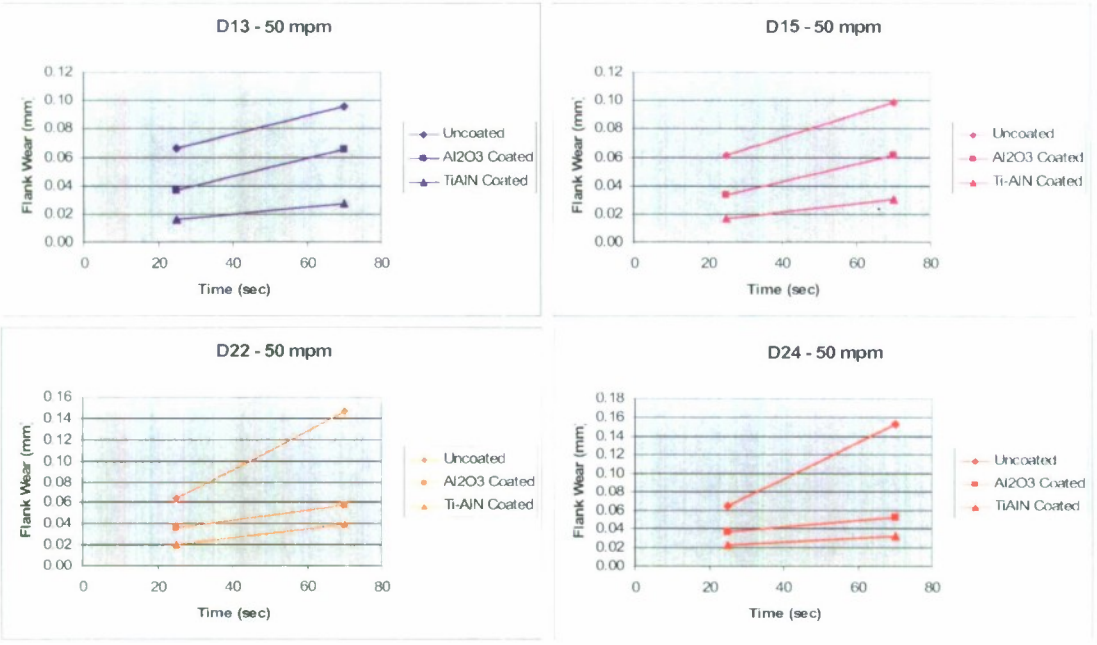


Figure 2 Wear Comparison for 4 alternative-binder tools, uncoated and coated

shows a significant effect due to binder proportion.

Figure 4 shows the results from a cutting speed of V = 100 mpm. Although no direct comparisons (i.e., head-to-head tests) were made at 100 mpm with uncoated tools, it is clear that the coated tools performed dramatically better. All of the best results obtained with uncoated tools run at V = 75 mpm still had major wear (VB greater than 0.3 mm) before 40 seconds of cutting. As can be seen in Figure 4, all of the coated tools were still cutting with minimal wear (VB less than 0.15 mm) after 75 seconds, and this at an even faster speed (V = 100 mpm). This suggests at least 2-3 times greater tool life with either coating. With respect to differentiation between grades, however, the tests run at 100 mpm showed no significant effect of tool grades; i.e., no effects due to binder proportion or presence of TiC/TaC. The overall repeatability was not as good with these tests as compared to the slower tests, possibly due to more random variation in the bar diameter, inconsistency in measuring a larger overall wear land with an irregular shape, and/or a fairly persistent occurrence of notch wear (at the depth of cut line) due to the higher speed which contributed to the irregular wear land.

Run-to-Failure Tests

A select set of additional tests were run to further investigate the effects and differences between the two types of coating. All of these additional tests were run on the same tool grade, d21. This grade has the same basic makeup as d24 but with a slightly different binder composition. d21 has 85% superalloy and 15% rhenium in the binder, while d24 has 60% superalloy, 25% cobalt, and 15% rhenium; i.e., more superalloy with d21, more cobalt with d24. These tests ran for longer than the short-time tests and were based on just a single cutting edge. After cutting was stopped and tool wear measured for each

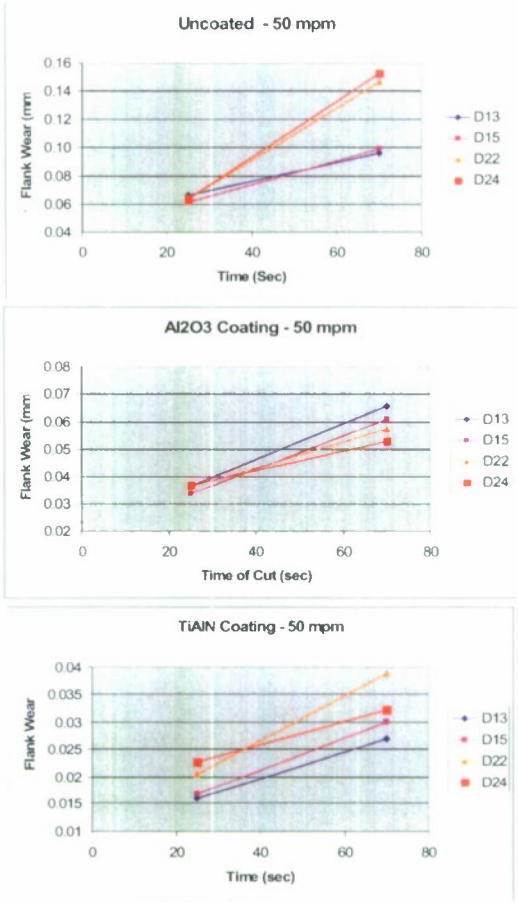


Figure 3 Comparison of 4 grades for each coating condition (V=50 mpm)

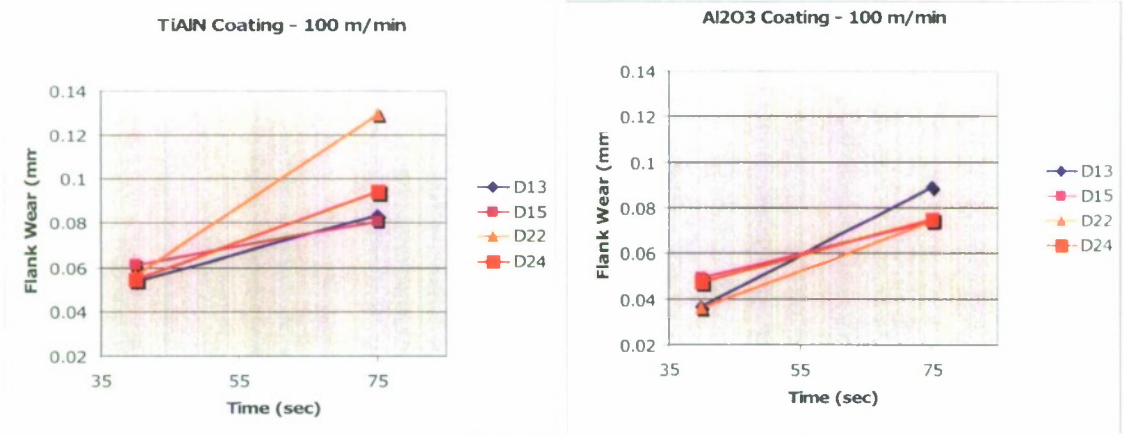


Figure 4 Comparison of 4 grades for the two coatings (V=100 mpm)

time, the same edge was placed back in the holder and cutting continued on that edge. The results from the first test of this type, run at V = 50 mpm, is shown in Figure 5. Although both coatings appear to perform about the same for the lower level of wear, after about VB = 0.1 mm the tool coated with Al_2O_3 performed somewhat better.

Although the d21 tool appears to be wearing a little faster (about VB = .07 mm after 70 seconds) than the four tools tested in the short-time tests (VB = .025 to .065 mm after 70 seconds), these eannot be eonsidered head-to-head comparisons sinee the bar diameter for these run-to-failure tests was signifieantly smaller resulting in more heat build-up and faster wear expeeted.

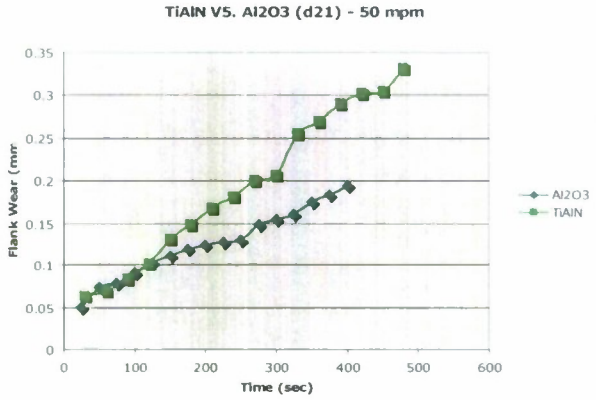


Figure 5 Coating Comparison at V = 50 mpm on d21 grade

Run-to-failure tests were also run at V = 100 mpm, with the results shown in Figure 6. The two graphs aetually show duplieated experiments run a couple of months apart. The graph on the left is based on just one eutting edge for eaeh curve (i.e., using edge replaeement) and shows little differenee between the eoatings. For comparison, an edge from the eommercially available eutting insert (Valenite VC912) was run with the same conditions. It appeared to do somewhat better than the alternative binder tools, but not signifieantly so. The graph on the right side of Figure 6 is based on replieated tests (i.e., results from two edges averaged for eaeh eurve), and the Al_2O_3 eoating providing a slight

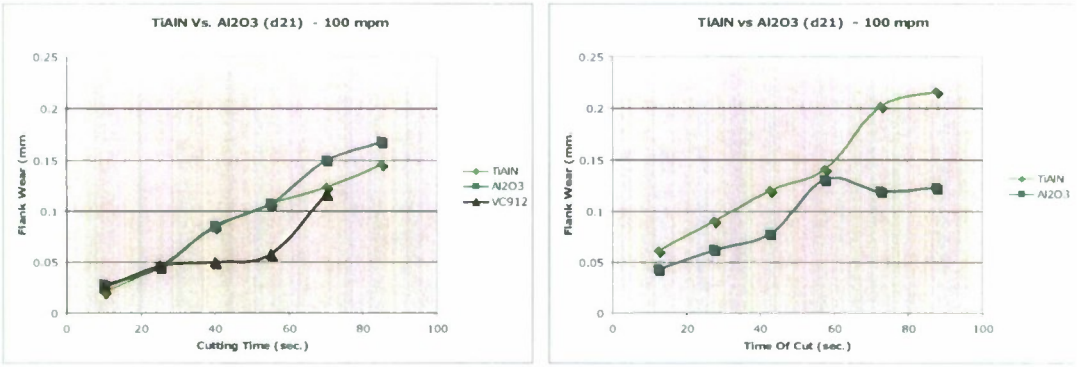


Figure 6 Coating Comparison at V = 100 mpm on d21 grade (duplicated tests)

advantage. Clearly, further experimentation is needed to confirm a preferenee for one eoating or another.

Conclusions

A new set of five grades of tungsten-carbide-based cutting tool samples has been produced using rhenium, cobalt, and nickel-based superalloy as binder metals, and two thin coatings were applied to subsets of each grade. The samples were used to cut Inconel 718 and the tool wear was measured and compared to uncoated tools and one commercially available tool. Much was learned about the effects of the coatings, though additional testing still has much to reveal.

Based on the testing completed for this project, the following conclusions can be drawn:

- Both the PVD TiAlN and the CVD Al_2O_3 coating provided tremendous protection for the tools. A dramatic increase in tool life (at least 2 to 3 times) resulting from the coatings is projected based on the short-time tests run for this study.
- Evidence suggests that the TiAlN coating performed better at the lower levels of wear (less than 0.1 mm wear land width) when cutting at the lower speed level (V = 50 mpm). When cutting for longer times or at the higher speeds, no conclusion could be drawn about the superior performance of either coating.
- Only minor evidence was found to suggest that the presence of TiC and TaC in the carbide phase of the tool composition had any effect on wear resistance, and even this evidence is contradictory.
- No evidence was found to support the conclusion that binder proportion has a significant effect on wear resistance.

Additional testing along the lines of the tests completed here are ongoing to investigate the potential for a more definitive answer on which coating is preferable and on whether binder proportion or the presence of TiC and TaC have a real effect. The PI and Dr. Liu were successful in securing both Phase I ($100,000) and Phase IB ($50,000) funding from the Small Business Innovative Research (SBIR) program from NSF to continue work on this project during the 2007-08 academic year. A research proposal and a business plan were recently submitted to NSF requesting Phase II funding (2 years, $500,000), with a notification expected in Spring 2008. With the launch of the new NSF-sponsored SBIR research, it appears that the researchers will get the chance to confirm and extend some the findings from this ONR-sponsored project.

Crack Diagnosis of a Rotor from Torsional/Lateral Coupling Mechanism

Projeet Investigators:

Xi Wu and James Meagher
Mechanical Engineering
California Polyteehnie State University
San Luis Obispo, CA

Project Overview

The purpose of this investigation into the vibration behavior of a rotor with transverse cracks is to determine whether and in which way a crack can be recognized by measuring the shaft vibrations so that catastrophic failures in rotors can be prevented and maintenance may be intelligently scheduled. Fatigue cracks/flaws are a potential source of catastrophic failures in rotating machinery. The sudden failures induced by fatigue cracks release considerable energy stored in the rotating system, sometimes causing a rotor to fly apart. Despite considerable efforts spent to develop a foolproof and reliable strategy to detect cracked rotors in aerospace, military, and power generation applications, current strategies remain inadequate. ONR Science and Technology Objectives: Logistics Log STO-2: Predictive maintenance systems and Maneuver MVR STO-1: Advanced power plants, drive trains, and suspensions relate to this study.

The long-term goal of the proposed research is to develop both analytical and experimental methods to quantitatively characterize the torsional/lateral responses of a two-mass cracked rotor. This study focuses on developing a nonlinear dynamic analytical model of a torsional/lateral coupled extended Jeffcott rotor with a crack, laterally loaded with a constant radial force and a rotating unbalance. This is considered to be an accurate yet tractable model that represents a wide class of machines. The contribution of this research is the development of a practical and realizable diagnostic method for rotating machinery which can predict the existence and severity of a cracked shaft from the vibration measurements typically available by high speed rotating machinery. The rotors of turbines and generators could be represented by a simplified equivalent configuration in which the masses of the turbine rotor and the generator rotor are concentrated on disks mounted on a light elastic shaft with two orthogonal lateral and two angular (torsional) degrees of freedom. Earlier work on the same topic simulated a crack as an asymmetric shaft, which is not accurate and cannot capture the whole profiles of a cracked rotor. Crack specific signatures predicted by simulation will be experimentally tested in future studies. As a result of mutual influence between lateral and torsional vibrations due to the crack, the vibration patterns, orbit shape and frequency spectra are unique and could be useful indicators for crack diagnosis. An expected outcome of the research is the development of more accurate and feasible nonlinear torsional/lateral coupled cracked model that may be used for automated machinery fault detection.

Major Findings and Results

- A practical cracked rotor model representing a wide class of machines has been developed.
- Results of this model demonstrate several diagnostic indicators which may be useful to detect the presence and severity of shaft cracks:
  - lateral-torsional coupling.
  - torsional critical speed frequencies at fixed non-integer ratios of lateral to torsional natural frequency.
  - the spectrum of the torsional vibration is shown to contain super- synchronous critical speeds related to the lateral natural frequency.
  - for cracked rotors that are torsionally rigid rotors and without lateral vibration, substantial vibrations occur at one-third and one-half the fundamental torsional resonance. With an increase in crack depth, the resonance peaks at 1×, 2× and 3× of the torsional vibration increase dramatically.
  - for more flexible rotors with a crack, strong coupling between lateral and torsional vibrations due to the crack and side load appear. Nonlinear lateral-torsional coupling from a crack also shifts the fundamental resonance peak in the torsional vibration response. The resonance peak frequencies shift depending on the ratio of the lateral to torsional natural frequencies with the peak responses occurring at fixed non-integer values of the lateral natural frequency for any particular frequency ratio.
- Leveraged this study to obtain hardware to make specialized measurements and increase our capabilities.
- Preliminary experimental measurements of lateral vibrations have been conducted in preparation for more extensive verification testing.
- Designed and built a machine allowing manufacture of cracked shafts for further study.
- Results of this study have been presented at several conferences. Results to be published in the International Journal of Rotating Machinery are attached:

"Investigation of Coupled Lateral and Torsional Vibrations of a Cracked Rotor under Radial Load," Wu, X., Meagher, J., Judd, C., *IMAC XXV (2007), Society for Experimental Mechanics*, Orlando, Florida USA, February 19 –22, 2007.

"Characterization and Modeling of Shaft Cracks and Rotating Asymmetries," Meagher J.M., Wu X., ISCORMA-4, Calgary, Alberta, Canada 27-30 August 2007.

"A Two-disk Extended Jeffcott Rotor Model Distinguishing A Shaft Crack from other Rotating Asymmetries," Wu, X., Meagher, J., International Journal of Rotating Machinery, accepted Jan. 2008.

"Nonlinear Contact Analysis of Gear Teeth for Malfunction Diagnostics," Kong D., Meagher J.M., Xu C., Wu X., Wu Y., *IMAC XXVI , Society for Experimental Mechanics*, Orlando, Florida USA, February 4 –7, 2008.

Background

A comprehensive literature survey of various crack modeling techniques and system behaviors of cracked rotors can be found in Wauer [1] and Dimarogonas [2]. More than 300 papers are cited. Diagnostic indicators of cracked rotors reported in these cited studies include: changes in lateral resonances due to flexibility changes in a cracked rotor, coupling that can be recognized from additional harmonics in the frequency spectrum, sub-harmonic resonances at approximately half and one third of the bending critical speed of the rotor [3-5]. Nonlinear modeling of the crack began with simple step functions that were later improved upon by Mayes. Mayes model [6] is more practical for deep cracks than a hinged model. Based on Mayes modified model, Sawicki and Wu et al. [7, 8, 9] studied the transient vibration response of a cracked Jeffcott rotor under constant acceleration ratios and under constant external torque. Muszynska et al. [10] and Bently et al. [11] discuss rotor coupled lateral and torsional vibrations due to unbalance, as well as due to shaft anisotropy under a constant radial preload force. Their experimental results exhibited the existence of significant torsional vibrations due to coupling with the lateral modes. In Bently and Muszynska's experiments, an asymmetric shaft was used to simulate the behavior of a crack. Our research extends and is extensively based upon the investigations of both Bently et al. [11] and Wu's work [12]. Starting from energy equations, an analytical model with four degrees of freedom for a torsional/lateral coupled rotor due to a crack is developed. A radial constant force is applied to the outboard disk to emphasize the effects of the gravity force which plays a critical role for the "breathing" of a crack. As preload increases, the vibration amplitudes in both lateral and torsional directions increase, making the measurements much easier to measure. The "second-order" nonlinear coupling terms due to a crack introduce super-synchronous peaks at certain rotational speeds, which is unique for a cracked rotor and might be used as an un-ambiguous crack indicator. Computer simulations also show that the rotational speeds at which amplitudes of the torsional vibrations reach maximum are governed by the ratio of lateral to torsional natural frequency.

Analytical Model

As reported in the 2005/06 year end report, an analytical model of an extended Jeffcott rotor, Figure 1, was used to derive equations of motion which take into consideration lateral/torsional vibration coupling mechanism induced by a "breathing" crack. Details of the modeling can be found in the 2005/06 report and in the literature published as a result of this study and listed above. The equation predict the lateral motion at the midspan (X,Y) and the torsional motion inboard (Θ) and outboard (θ).

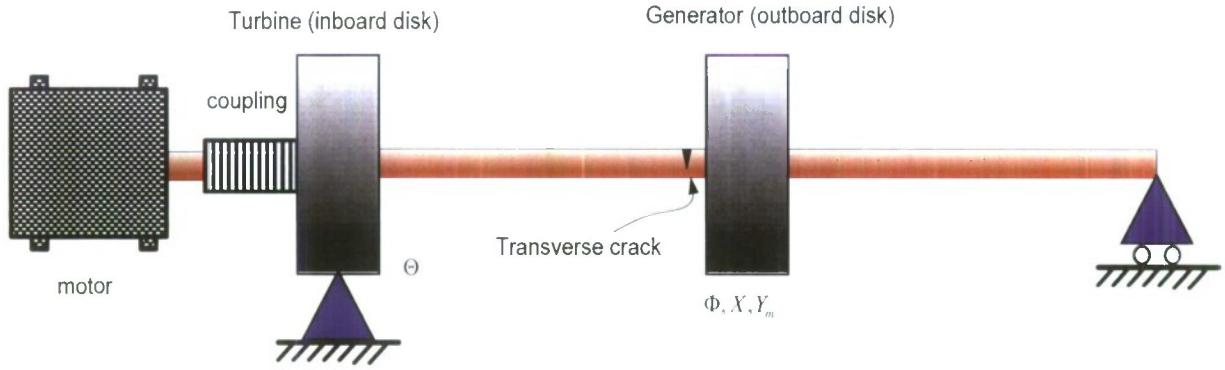


Figure 1. Configuration of the cracked extended Jeffcott rotor with two disks.

General equations of motion obtained using Lagrange's method:

$$\ddot{X}+\frac{C}{M}\dot{X}+\omega_n^2\left[1-\frac{f(\Phi)}{2}\left(\Delta k_1+\Delta k_2\cos 2\Phi\right)\right]X-\frac{\omega_n^2 f(\Phi)\Delta k_2\sin 2\Phi}{2}\left(Y_m-\frac{P}{K}\right)$$
$$=\varepsilon\left(\Omega+\dot{\varphi}\right)^2\cos\left(\Phi+\delta\right)+\varepsilon\ddot{\varphi}\sin\left(\Phi+\delta\right)$$

$$\ddot{Y}_m+\frac{C}{M}\dot{Y}_m-\frac{\omega_n^2 f(\Phi)\Delta k_2\sin 2\Phi}{2}X+\omega_n^2\left[1-\frac{f(\Phi)}{2}\left(\Delta k_1-\Delta k_2\cos 2\Phi\right)\right]Y_m$$
$$=\varepsilon\left(\Omega+\dot{\varphi}\right)^2\sin\left(\Phi+\delta\right)-\varepsilon\ddot{\varphi}\cos\left(\Phi+\delta\right)-\frac{Pf(\Phi)}{2M}\left(\Delta k_1-\Delta k_2\cos 2\Phi\right)$$

$$\ddot{\theta}+\frac{K_t+K_c}{I_0}\theta-\frac{K_t}{I_0}\varphi=-\frac{C_t+C_c}{I_0}\dot{\theta}+\frac{C_t}{I_0}\dot{\varphi}$$

$$\ddot{\varphi}+\frac{C_t}{I}\dot{\varphi}-\frac{C_t}{I}\dot{\theta}+\frac{K_t}{I}\varphi-\frac{K_t}{I}\theta=\frac{\mathrm{P}\varepsilon f(\Phi)}{2I}\left(\Delta k_1\cos\left(\Phi+\delta\right)-\Delta k_2\cos\left(\Phi-\delta\right)\right)$$
$$+\frac{P^2}{2KI}\left[\frac{1}{2}\frac{\partial f(\Phi)}{\partial\Phi}\left(\Delta k_1-\Delta k_2\cos 2\Phi\right)+f(\Phi)\Delta k_2\sin 2\Phi\right]+\Gamma_c$$

$$\Gamma_c=-\frac{C\varepsilon}{\rho^2 M}\left[\dot{X}\sin\left(\Phi+\delta\right)-\dot{Y}_m\cos\left(\Phi+\delta\right)\right]+\frac{\varepsilon\omega_n^2}{\rho^2}\left(1-\frac{f(\Phi)}{2}\Delta k_1\right)\left[-X\sin\left(\Phi+\delta\right)+Y_m\cos\left(\Phi+\delta\right)\right]$$
$$+\frac{\varepsilon f(\Phi)\Delta k_2\omega_n^2}{2\rho^2}\left[-X\sin\left(\Phi-\delta\right)+Y_m\cos\left(\Phi-\delta\right)\right]$$
$$+\frac{X^2\omega_n^2}{2\rho^2}\left[\frac{1}{2}\frac{\partial f(\Phi)}{\partial\Phi}\left(\Delta k_1+\Delta k_2\cos 2\Phi\right)-f(\Phi)\Delta k_2\sin 2\Phi\right]$$
$$+\frac{Y_m\left(Y_m-\frac{2P}{K}\right)\omega_n^2}{2\rho^2}\left[\frac{1}{2}\frac{\partial f(\Phi)}{\partial\Phi}\left(\Delta k_1-\Delta k_2\cos 2\Phi\right)+f(\Phi)\Delta k_2\sin 2\Phi\right]$$
$$+\frac{\Delta k_2 X\left(Y_m-\frac{P}{K}\right)\omega_n^2}{2\rho^2}\left[\frac{\partial f(\Phi)}{\partial\Phi}\sin 2\Phi+2f(\Phi)\cos 2\Phi\right]$$

Simulation Results

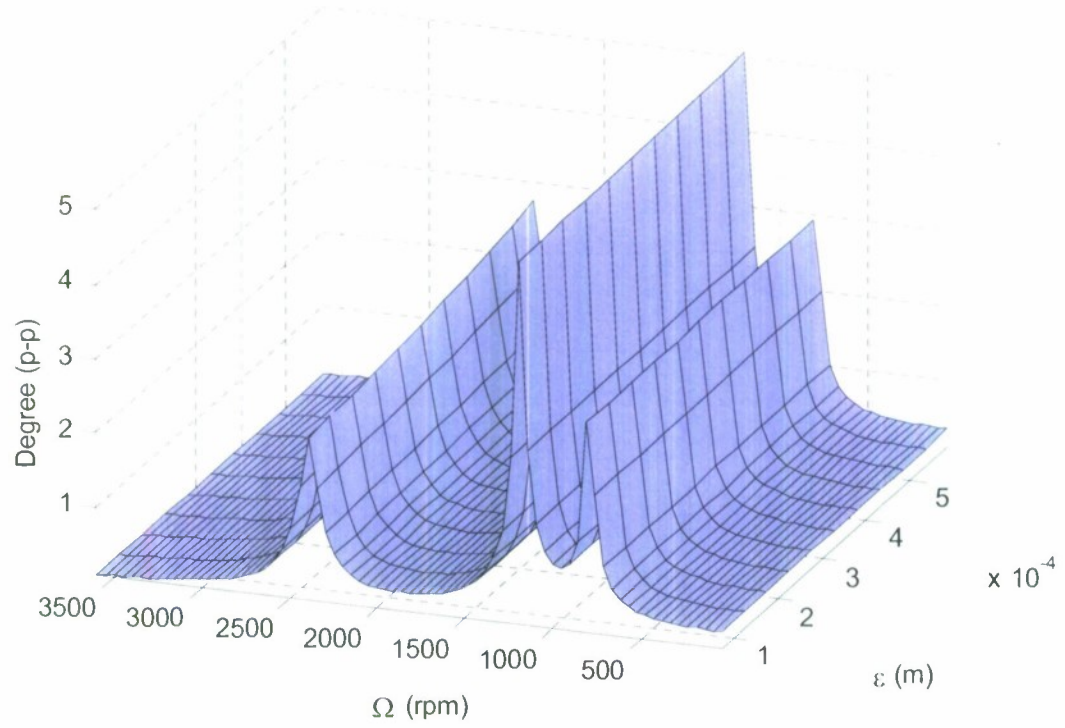
Figure 2 depicts the vibration response for a range of eccentricities for a cracked rotor and an asymmetric but uncracked rotor. For the cracked shaft and the asymmetric shaft, the critical speed associated with the 1x torsional natural frequency at 2400 rpm has a magnitude that increases with increasing eccentricity. The response at the other critical speeds is independent of eccentricity. The consequence of this is that for well balanced shafts the presence of a crack will be more easily detected by monitoring the response at a shaft rotative speed of $\omega_t/2$ or $\omega_t/3$. At large values of eccentricity the response at shaft rotative speeds equal to the torsional natural frequency dominate. The sensitivity of response to changing eccentricity is much greater for the asymmetric shaft.

The frequency response at various shaft rotative speeds, Ω, for the same system is depicted in Figure 3. Each is dominated by 2x responses with an additional 3x order response for the cracked shaft. The 2x response has typically been used to monitor the presence of cracks but this shows that the response is obscured by asymmetry not associated with a malfunction. Numerous other simulation results that support the major findings listed above can be found in references [13-15].

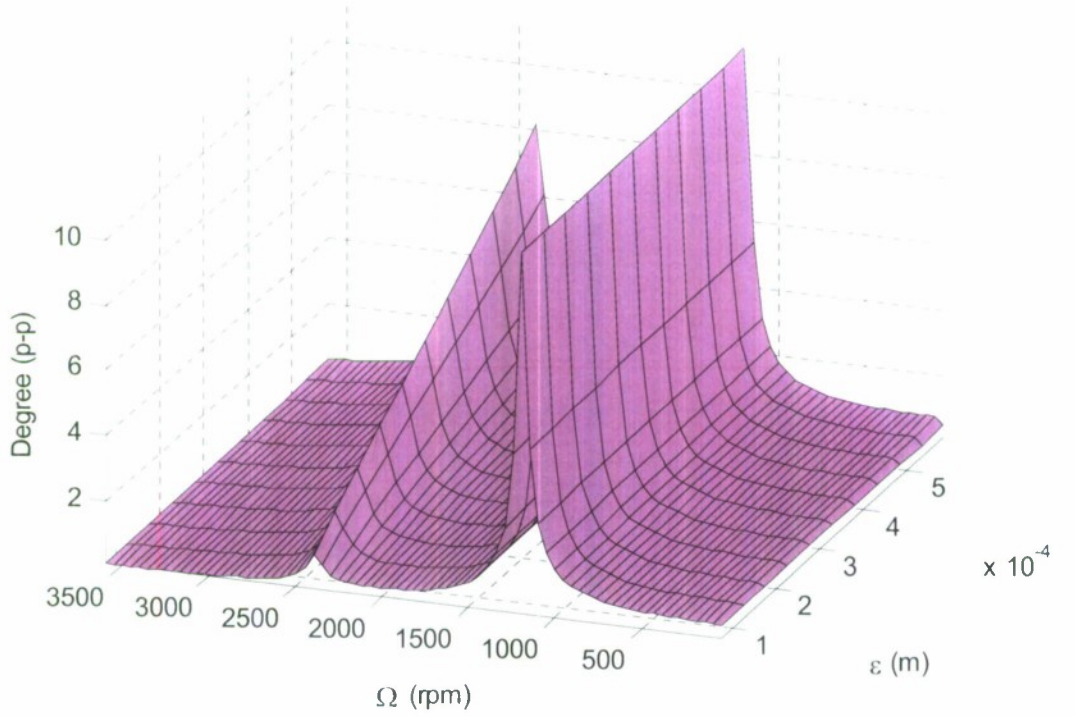
Current Work

Based upon suggestions from ONR reviewers, we extended our studies to gear box malfunctions. Shown below in figure 4 is a gear box modeled with MSC-ADAMS software. This is a model of an actual gear box we have installed into our laboratory and instrumented with vibration probes. An efficient and accurate method to model gear tooth dynamics using modern engineering tools is currently being developed. A CAD model is transferred into multi-body dynamics software. A composite rigid body dynamics and contact model developed for the tooth loads is possible using this software. The method allows for nonlinear contact mechanics to include the effects of friction, damping, and hertzian-contact in a localized region of gear mesh. The localization of the elastic portion of the model apart from the rigid body portions reduces computing time.

Simulation results such as shown in figure 5 indicate that this method may be useful to predict the fault pattern of the gearbox. Presently we are refining this model, developing FEA models for verification purposes, and instrumenting the gear box. As data becomes available we will be able to bench mark our model and determine if gear malfunction frequencies are discernable.

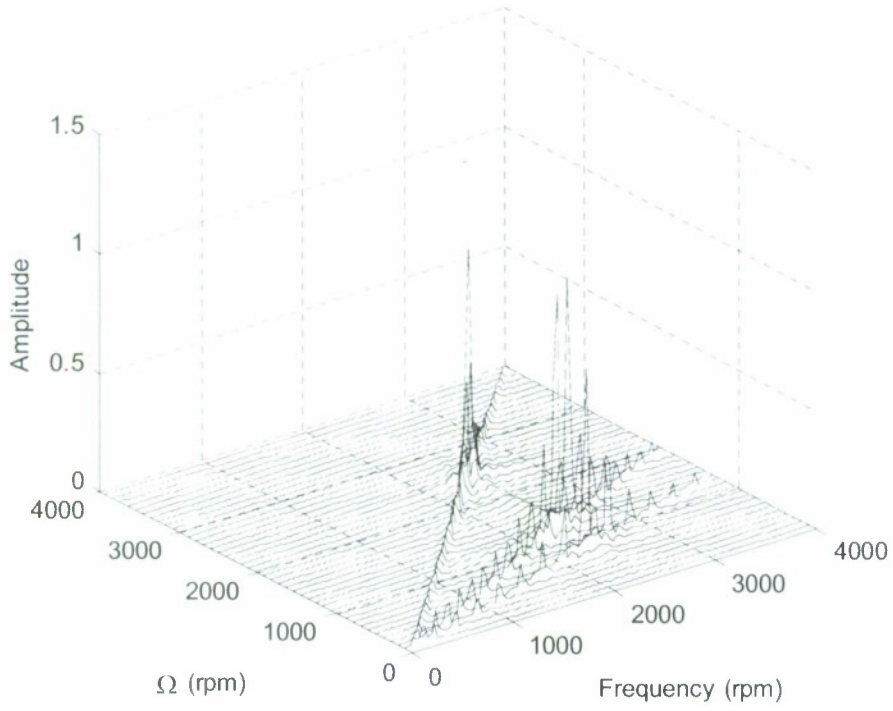


(a). $\Delta k_{\xi} / K = 0.30$

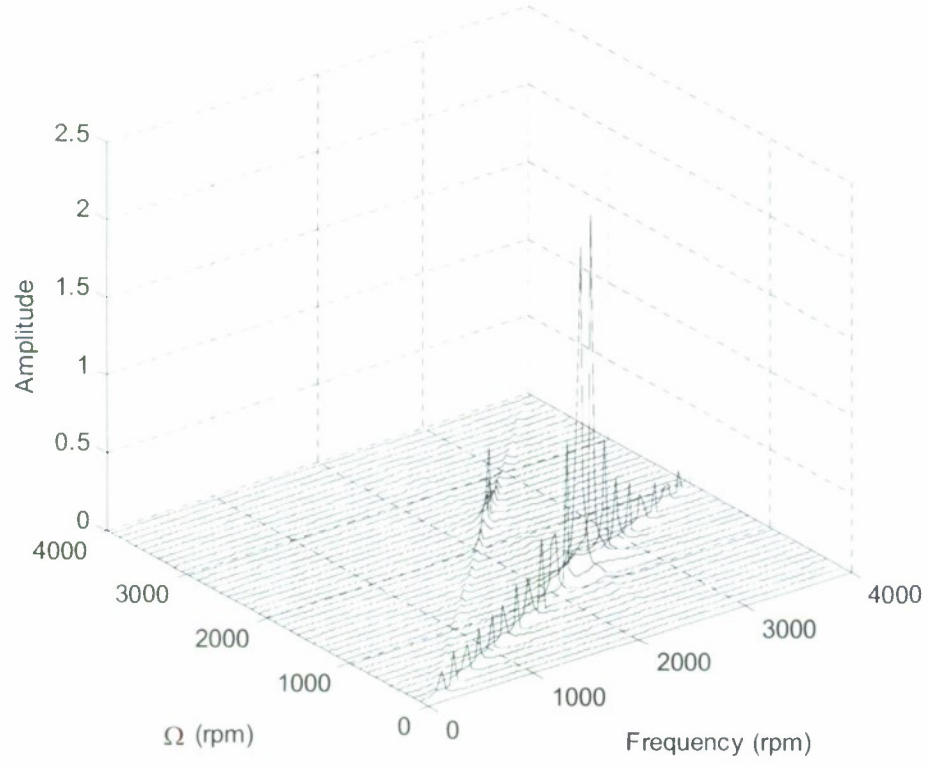


(b). q = 0.14

Figure 2. Overall peak-to-peak torsional vibration response, φ, Special Cases, for cracked shaft (a) and asymmetric shaft (b).



(a). $\Delta k_{\xi} / K = 0.30$



(b). q = 0.14

Figure 3. Half spectrum of zero-to-peak torsional vibration response, φ, Special Cases, for cracked shaft (a) and asymmetric shaft (b).

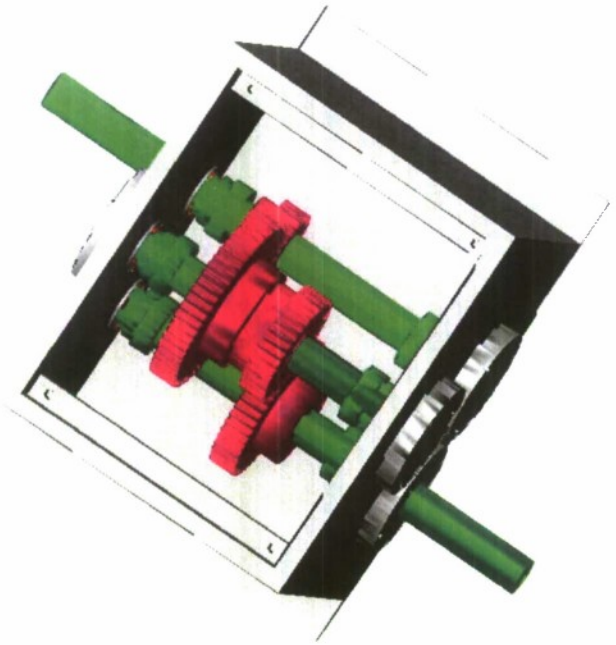


Figure 4: MSC-ADAMS gearbox model

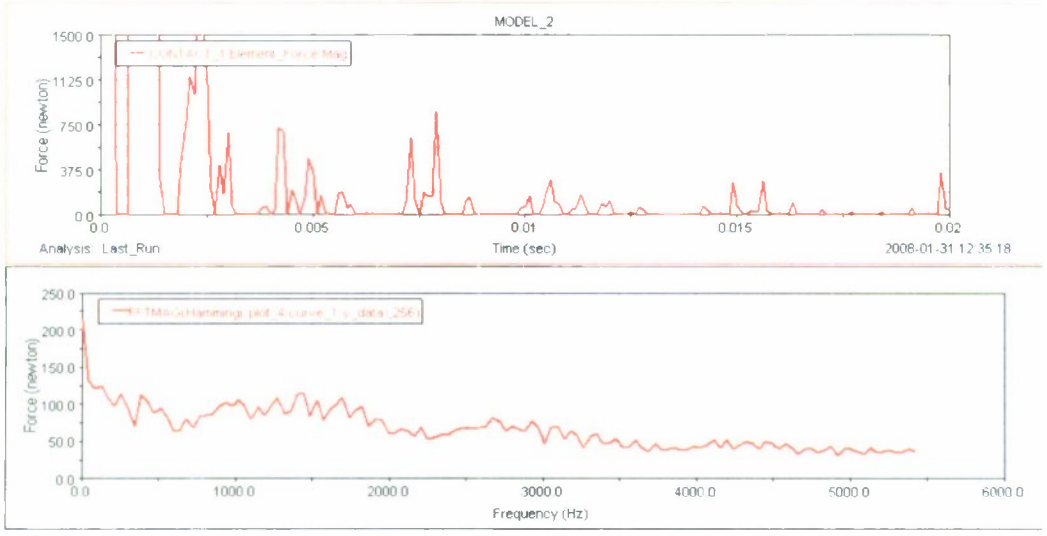


Figure 5: Contact force time history and FFT spectrum between the 1st meshing gears

Nomenclature

ω_n, ω_t	lateral and torsional natural frequencies, respectively.
Θ	angular displacement of inboard disk.
θ	angular rotation of the inboard disk relative to motor.
θ_0	initial angular location of inboard disk.
Φ	angular rotation of outboard disk.
φ	angular rotation of the outboard disk relative to motor.
φ_0	initial angular location of outboard disk.
I_0	inboard disk polar moment of inertia.
I	outboard disk polar moment of inertia.
δ	angular orientation of outboard disk eccentricity.
Γ_c	lateral coupling terms in outboard disk equation of motion.
M	outboard disk mass.
P	vertical side load.
X,Y	outboard disk lateral motion in inertial coordinates.
ξ,η	rotor fixed rotating coordinates.
Y_m	dynamic vertical vibration in inertial coordinates.
ε	eccentricity of outboard disk.
C_c, K_c	motor-shaft coupling damping and stiffness, respectively.
C, C_t	lateral and torsional damping coefficients, respectively.
ζ, ζ_t	lateral and torsional damping ratios, respectively.
Ω	motor speed.
K_t	torsional shaft stiffness.
K	uncracked shaft lateral stiffness.
$f(\Phi)$	crack steering function.
ρ :	radius of gyration.
Δk_ξ, Δk_η	reduced stiffness in ξ and η directions, respectively.

References

[1]. Wauer, J., 1990, "On the Dynamics of Cracked Rotors: A Literature Survey", *Applied Mechanics Reviews*, Vol. 43(1), pp. 13-17.

[2]. Dimarogonas, A. D., 1996, "Vibration of Cracked Structures: A State of the Art Review," *Engineering Fracture Mechanics*, Vol. 55 (5), pp. 831 – 857.

[3]. Gasch, R., 1976, "Dynamic behavior of a simple rotor with a cross-sectional crack", *Paper C178/76, I. Mech. E. Conference on Vibrations in Rotating Machinery*, pp.123-128.

[4]. Gasch, R. A, 1993, "Survey of the Dynamic Behavior of a Simple Rotating Shaft with a Transverse Crack", *Journal of Sound and Vibration* Vol. 162, pp313-332.

[5]. Chan, R. K. C. and Lai, T. C., 1995, "Digital Simulation of a rotating shaft with a transverse Crack," *Appl. Math. Modelling* Vol. 19, July pp. 411-420.

[6]. Mayes, I. W. and Davies, W. G. R., 1984, "Analysis of the Response of a Multi-Rotor-Bearing System Containing a Transverse Crack in a Rotor", *ASME Journal of Vibration, Acoustics, Stress, and Reliability in Design*, Vol. 106, pp 139-145.

[7]. Sawicki, J. T., Wu, X., Baaklini, G.Y. and Gyekenyesi, A., 2003, "Vibration-Based Crack Diagnosis in Rotating Shafts During Acceleration Through Resonance", *Proceedings of SPIE*, Vol. 5046.

[8]. Sawicki, J. T., Bently, D. E., Wu, X., Baaklini, G.Y. and Friswell, M. I., 2003, "Dynamic Behavior of Cracked Flexible Rotor Subjected to Constant Driving Torque", ISCORMA-2, Gdańsk, Poland, 4-8, August 2003, pp. 231-241.

[9]. Sawicki, J.T., Wu, X., Gyekenyesi, A. L. and Baaklini, G. Y., 2005 "Application of Nonlinear Dynamic Analysis for Diagnosis of Cracked Rotor Vibration Signatures", submitted to SPIE International Symposium, San Diego, California USA, March 7-10.

[10]. Muszynska, A., Goldman, P. and Bently, D. E., 1992, "Torsional/Lateral Cross-Coupled Responses Due to Shaft Anisotropy: A New Tool in Shaft Crack Detection", *I. Mech. E., C432-090, Bath, United Kingdom*, pp. 257-262.

[11]. Bently, D. E., Goldman, P. and Muszynska, A., 1997, ""Snapping" Torsional Response of an Anisotropic Radially Loaded Rotor," *Journal of Engineering for Gas Turbines and Power*, Vol. 119, pp. 397-403.

[12]. Wu, X., 2005, Doctoral dissertation, "Vibration-based Crack-induced Damage Detection of Shaft-disk System", Cleveland State University.

[13]. Wu, X., Meagher, J., Judd, C., "Investigation of Coupled Lateral and Torsional Vibrations of a Cracked Rotor under Radial Load,", *IMAC XXV (2007), Society for Experimental Mechanics*, Orlando, Florida USA, February 19 –22, 2007.

[14]. Meagher J.M., Wu X., "Characterization and Modeling of Shaft Cracks and Rotating Asymmetries," ISCORMA-4, Calgary, Alberta, Canada 27-30 August 2007.

[15]. Wu, X., Meagher, J. "A Two-disk Extended Jeffcott Rotor Model Distinguishing A Shaft Crack from other Rotating Asymmetries," International Journal of Rotating Machinery, accepted Jan. 2008.

[16]. Kong D., Meagher J.M., Xu C., Wu X., Wu Y. ,"Nonlinear Contact Analysis of Gear Teeth for Malfunction Diagnostics,", *IMAC XXVI , Society for Experimental Mechanics*, Orlando, Florida USA, February 4 –7, 2008.

Implementation of a Neural Network Controller for Real-Time Adaptive Control

Project Investigator:

Xiao-Hua (Helen) Yu
Electrical Engineering
California Polytechnic State University
San Luis Obispo, CA

Abstract

Technical Merit:

In this research, a self-tuning controller based on multi-layer feed-forward neural network is developed for real-time output voltage regulation of a class of DC power supplies. The neural network based controller has the advantage of adaptive learning ability, and can work under the situations when the input voltage and load current fluctuate. Levenberg-Marquardt back-propagation training algorithm is used in computer simulation. The neural network controller is implemented and tested on hardware using a DSP (digital signal processor). Experimental results show that this neural network based approach outperforms the conventional analog controller, in terms of both line regulation and load regulation.

Publications Resulting from the Project:

Two technical papers (both co-authored with EE graduate student under my supervision) are published and presented at international conferences this year, including the IEEE international conference on control and automation (May 2007), and the IEEE international conference on neural networks (August 2007).

1. Introduction

DC-DC converters can be found in almost every electronic device nowadays, since all semiconductor components are powered by DC source. PWM (pulse-width modulation) technique is used in many voltage regulators. It changes the average value (i.e., dc component) of a square waveform by modulating its duty cycle.

One of the design objectives for electronic engineers is to improve the efficiency of power conversion. In PWM converters, a switching network is employed for square waveform modulation. Ideally, the power dissipated by the switch network is zero; however in practice, the power efficiency of a typical DC-DC converter could be as low as 70%. Many different kinds of topologies ([1], [2], [3]) have been investigated in the past to reduce the switching loss. Unfortunately, those topologies either need additional components for the power circuit, which may introduce some unstable factors to the circuit; or operate at variable frequency, which makes the filter design at output stage very difficult. Phase-shifted zero-voltage switching full-bridge converters overcome the above problems and thus have been received more and more attention recently. It employs zero-voltage-switching (ZVS) technique which allows the voltages across the transistors to swing to zero just before the start of the next conduction cycle ([4], [5], [6]).

The conventional approach assumes that the circuit is operated around its equilibrium state, and then derives a set of linear equations based on this assumption [9] [10]. However, in practice, the supply voltage and load current may have a wide range of variation; so the controller has to be designed to work under such conditions.

Artificial neural networks (ANN) have been widely used in the field of system identification, adaptive control, and statistic modeling in recent years. A neural network is

composed of many non-linear adaptive processing elements and is capable of approximating any measurable function under certain conditions.

Recently, artificial neural network based controller has been chosen as an alternative to classic methods ([7], [8]) to improve the performance of DC power supply to dynamical system changes, due to the adaptive learning ability of the neural network controller. However, no prior work has been done to control a PSFB (Phase-Shifted Full-Bridge) converter using the neural network approach yet. In this research, an approach based on multi-layer feed-forward neural network controller is investigated. A Matlab Simulink model is developed first to generate the data set; then the neural network is trained by Levenberg-Marquardt back-propagation algorithm. Finally, the neural controller is implemented on a DSP evaluation board eZdsp® F2812 (with digital processor TMS320F2812) and a phase-shifted zero-voltage-switching circuit board UCC3895EVM. Satisfactory experimental results with the neural network controller are obtained and compared with the conventional analog controller, in terms of both line regulation and load regulation.

2. Experimental Results

The overall system block diagram is shown in Fig. 1. The inputs to the neural network include input voltage, load current, and the change of output voltage. The control objective is to keep the output voltage stable (at the nominal value of 3.3V in this case) under the conditions of different input voltages and load currents.

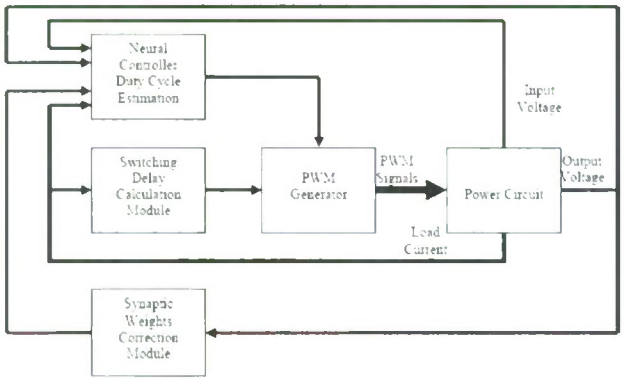


Fig. 1. The System Block Diagram

In order to investigate the performance of the neural network controller from experimental data, a DSP (Digital Signal Processor) evaluation module eZdsp® F2812 and a phase-shifted zero-voltage-switching evaluation board UCC3895EVM are used to implement the neural controller and the power circuit, respectively. The on-board digital signal processor TMS320F2812 is a 32-bit CPU with 150 MIPS (million instructions per second) operating speed. The phase-shifted full-bridge converter is operated on 48V (nominal input) and can provide an output of 3.3V. The neural controller is developed in C language. As shown in Fig. 2, the PWM generator and the power circuit blocks are implemented on board UCC3895EVM and the rest of the blocks (i.e., the neural network controller and the switching delay calculation module) are implemented with DSP TMS320F2812.

First, the analog signals from the power circuit are digitized and processed before being fed into the neural network's inputs. Since the power circuit is driven by 400 kHz high frequency PWM signals, the raw output voltage signal is quite noisy during the switch turn-on and turn-off

time. To solve this problem, multiple samples are taken to obtain the average value over a certain period of time. Next, the processed data is fed into the neural network to estimate the desired duty cycle. The new estimated duty cycle is then applied to the power circuit, and the output voltage is monitored. The computational flow chart is shown in Fig. 3.

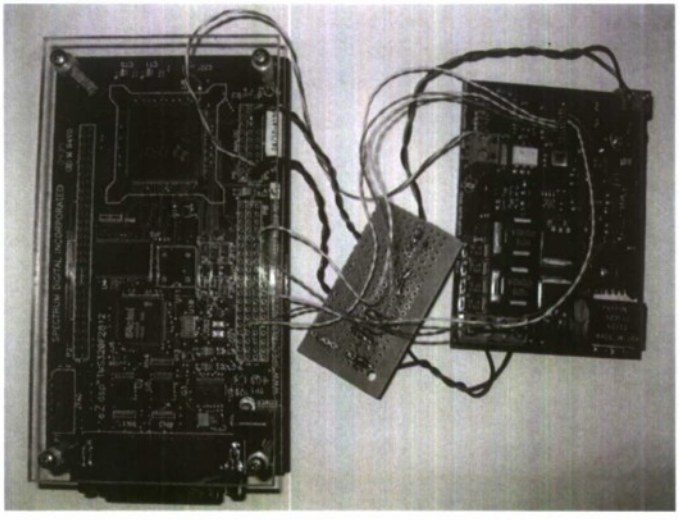


Fig. 2. Hardware Implementation

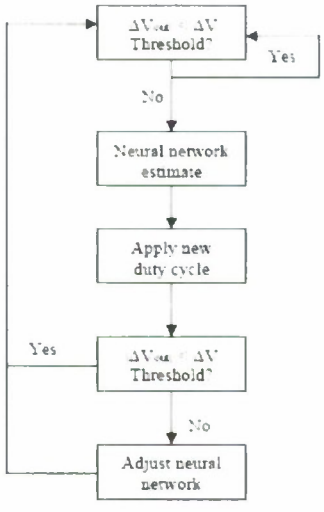


Fig. 3. The Flow Chart

At nominal input voltage 48V, the converter's steady-state output voltage is measured under different load conditions and the result is shown in Figure 4, where the curve on top is the output controlled by neural network controller and the other one is the output controlled by a conventional analog controller (UCC3895). The desired nominal output voltage is 3.3V by design. As shown in the plot, the maximum deviation on output by the neural network controller is 0.08V, which results in about 2.4% of error; while the maximum deviation on output by the analog controller is 0.1V, which results in about 3.0% of error. Obviously, the overall neural network controller outperforms the conventional analog controller.

The percentage of load regulation gives us an indication of how much the output voltage changes over a range of load resistance values. In constant voltage mode, variations in the load result in changes in the load current. Experimental results show that the percentage of load regulation using the neural network controller is only about 2.1%, which concludes that the neural controller provides excellent load regulation (comparing to the load regulation of 4.5% by the conventional controller).

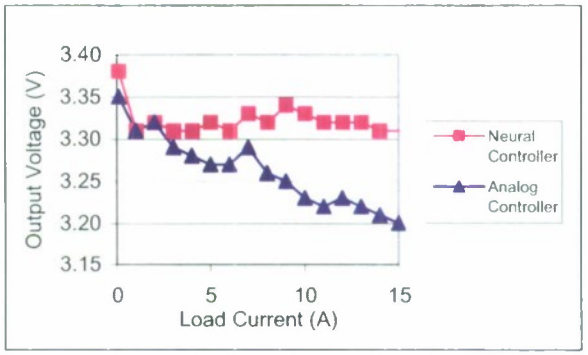


Fig. 4. Load Regulation at 48V Nominal Input Voltage

At full load condition, the converter's output voltage is measured at different input voltages, and the result is shown in Fig. 5. The maximum output error is only 0.04V, which is about 1.2% of the nominal value (the error rate of the analog controller is about 3.0%). Another performance measure, the percentage of line regulation between 48V and 72V input voltage is also calculated and compared. The neural controller achieves better line regulation (0.9%) than the analog controller (3.9%).

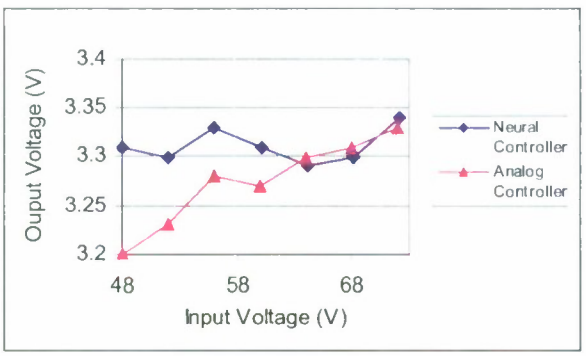


Fig. 5. Line Regulation at Full Load 15A

3. Conclusion

A neural controller as an alternative to classic controller for PSFB DC-DC converter is proposed in this research. Both simulation and experimental results show that the neural network controller is able to estimate the desired duty cycle under several dynamic conditions. In terms of both line regulation and load regulation, the neural network based approach outperforms the conventional analog controller. More research works will be done to optimize the software to speed up the neural controller for real-time applications.

References

[1] P. R. Chetty, "Resonant power supplies: Their history and status," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 7, no. 4, pp. 23–29, Apr. 1992.

[2] M. G. Kim and M. J. Youn, "An energy feedback control of series resonant converters," *IEEE Trans. Power Electron.*, vol. 6, no. 4, pp. 338–345, Jul. 1991.

[3] J. M. Carrasco, E. Galván, G. E. Valderrama, R. Ortega, and A. Stankovic, "Analysis and experimentation of nonlinear adaptive controllers for the series resonant converter," *IEEE Trans. Power Electron.*, vol. 15, no. 3, pp. 536–544, May 2000.

[4] H.-S. Choi and B. H. Cho, "Novel zero-current-switching (ZCS) PWM switch cell minimizing additional conduction loss," *IEEE Trans. On Industrial Electronics*, Vol. 49, No. 1, P. 165-172, February 2002.

[5] H.-S. Choi, J.-W. Kim, and B. H. Cho, "Novel zero-voltage and zero-current-switching (ZVZCS) full-bridge PWM converter using coupled output inductor," *IEEE Trans. On Power Electronics*, Vol. 17, No. 5, P. 641-648, September 2002.

[6] X. Ruan and Y. Yan, "A novel zero-voltage and zero-current-switching PWM full-bridge converter using two diodes in series with the lagging leg," *IEEE Trans. On Industrial Electronics*, Vol. 48, No. 4, P. 777-785, August 2001.

[7] J. M. Quero, J. M. Carrasco, and L. G. Franquelo, "Implementation of a neural controller for the series resonant converter," *IEEE Trans. On Industrial Electronics*, Vol. 49, No. 3, P. 628-639, June 2002.

[8] F. Kamran, R. G. Harley, B. Burton, T. G. Habetler, and M. A. Brooke, "A fast on-line neural-network training algorithm for a rectifier regulator," *IEEE Trans. On Power Electronics*, Vol. 13, No. 2, March 1998.

[9] B. Choi, J. Kim, B. H. Cho, S. Choi, and C. M. Wildrick, "Designing control loop for DC-to-DC converters loaded with unknown AC dynamics," *IEEE Trans. On Industrial Electronics*, Vol. 49, No. 4, P. 925-932, August 2002.

[10] Texas Instrument Inc, "BiCMOS advaneed phase shift PWM controller," Data sheet of UCC3895, January 2001.

[11] B. Andreyeak, "Designing a Phase Shifted Zero Voltage Transition (ZVT) Power Converter," High Frequency Power Conference Proceedings, 1992

Tracking Lips in Unconstrained Imagery for Automatic Speech Reading

Project Investigator:

Xiaozheng (Jane) Zhang
Electrical Engineering
California Polytechnic State University
San Luis Obispo, CA

1. Background

The primary goal of this projeet is to develop a robust lip identification and tracking system that is flexible enough to handle variations in lighting, baekground, and human speakers. This serves as an essential first step towards building a practieal audio-visual speeeh reeognition system. For the purpose of algorithm development and evaluation, we ineorporated AVICAR database [1] that eonsists of audio-visual data of 100 speakers speaking inside a car. The data was reeorded under five different noise conditions (car in idling, ear running at 35/55 mph with all windows rolled up/down), thus presents challenges for traeking lips under varying real-world test conditions. We also included data sets collected by us in the algorithm development.

Two graduate students (Brandon Crow and Stephen Kam) were supported by this grant and have performed under my supervision most of the work presented here.

The major findings ean be divided into two areas. One is the continued study of color model in lip representation. The seeond area is the design of the lip reader that traeks lips continuously in a sequence of image frames. In the following, we present detailed findings in eaeh area.

2. Our Approach

2.1 Color Model

In our previous work we justified the use of hue eolor in representing the lips because it offers the best separation between the lip region and the baekground pixel values based on the histograms obtained from hundreds of manually segmented lip images. In a subsequent study we further studied the invariant property of hue eolor with respect to varying shades of lipstiek and under different lighting eonditions. Speeifieally, we consider 1. Light Conditions: fluoreseents, sunlight, incandeseent; 2. Lip Color: black lipstick, maroon lipstiek, no lipstick (Natural), pale lipstick, bright Red lipstiek with glitter. We found that hue eolor is surprisingly eonstant under all varying eonditions. Figure 1 shows the normalized hue histograms for varying shades of lipstick. These results further eonfirm the usage of hue eolor in representing the lip pixels under varying conditions. It was found that skin hue has a normal distribution bounded (based on ensemble averaging) approximately by 0.8 and 0.2 (290° and 70°) – red hue. To use lip hue in a eonneeted region, in the following, a shift of 0.2 is applied to all hue values within the system [2], thereby shifting skin's hue distribution to 0 and 0.4 (0° and 140°). This HSI variant will henceforth be referred to as shifted-HSI (sHSI).

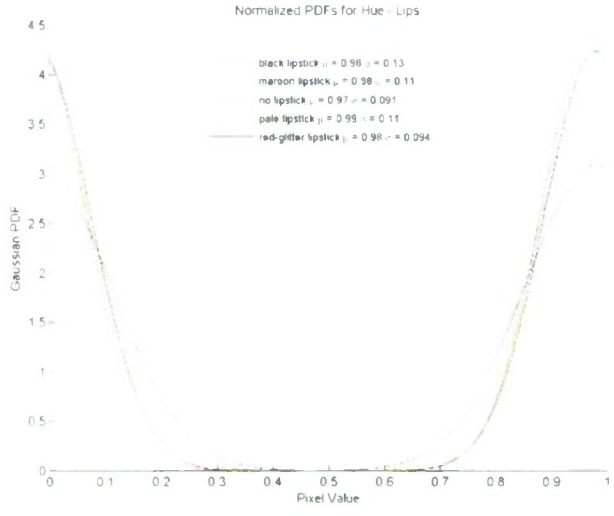


Figure 1 Hue Invariance Study for Varying Shades of Lipstick.

2.2 Lip Reader System

2.2.1 Find Face

This module attempts to locate a face within the current video frame. This is accomplished by selecting 5 regions within the frame and comparing the estimated probability distribution functions to that of three stored generic face models. The comparison is made using the Bhattacharyya coefficient, which provides rotational and scale invariance. A large Bhattacharyya coefficient represents a high degree of similarity between the stored face model and the target region. Therefore, noting that the target candidate may not contain the complete face (partial face, partial background or occlusion) an experimentally determined threshold of 0.6 is used. Specifically, if the Bhattacharyya coefficient for a target region and a face model is equal to or greater than 0.6, the system determines that a face has been found. Figure 1 depicts a top-level design of the *Find Face* module and a more in depth explanation of the module's functionality and results follow.

Generic Face Models

There are a total of three generic face models – representing light, medium and dark skin tones. To facilitate model/target comparison using the Bhattacharyya coefficient, both model and target are converted to statistical representations using histograms (pdf approximation) and the Epanechnikov kernel [2, 3]. The Epanechnikov kernel weights the image's pixels heavier toward the center and less at its perimeter, creating robustness towards edge and background noise. The histogram size used for the system is 16x16 (16 sH bins and 16 S bins). Figure 2 displays the three face models, their sHSI representations and their resulting pdf approximations. All video sequences were generated from a Channel Plus 7120 ¼" CCD camera or obtained from the AVICAR database [1].

Note: For clarity, the pdfs were scaled by 1/max[pdf] as their values are too low to see otherwise. Additionally, the center images are in sHSI color space but are

displayed in RGB. Therefore, red is shifted-hue, green is saturation, and blue is intensity.

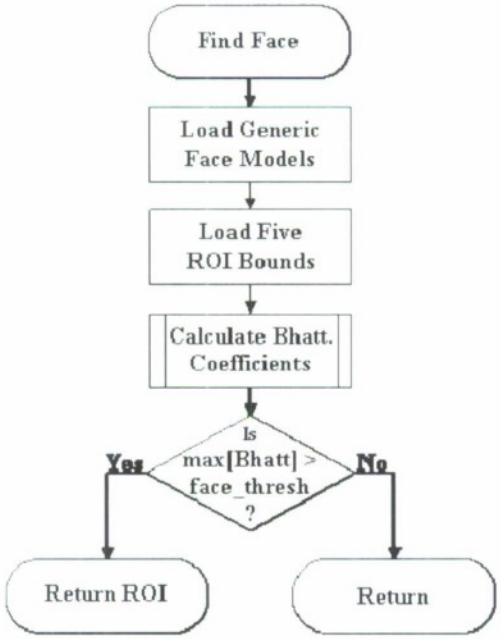


Figure 1: Top-Level Design of Find Face Module

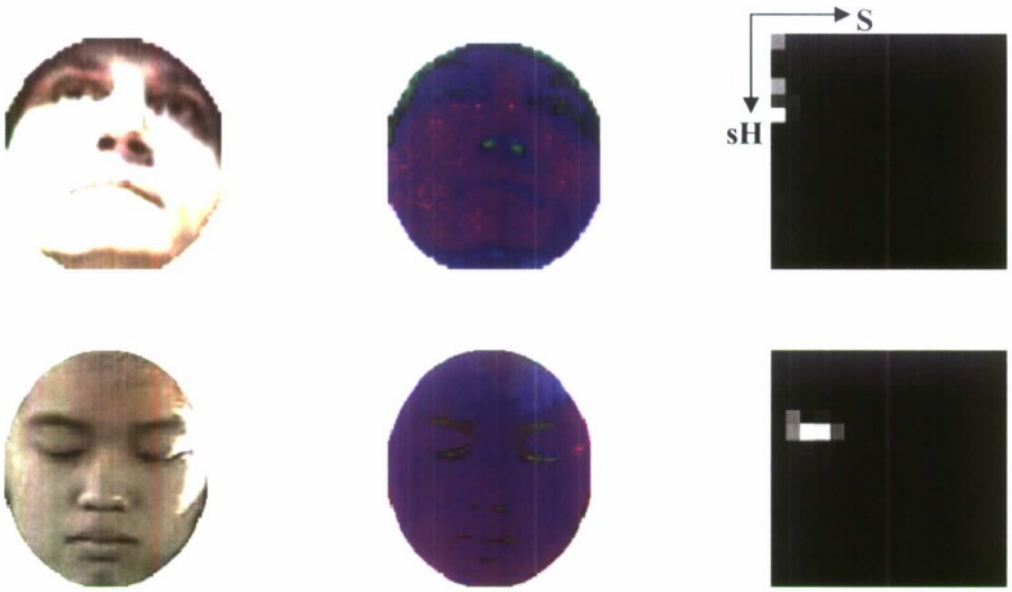


Figure 2: Face Model Conversion

Five ROI Bounds

The bounds for five regions of interest are constructed to analyze the four corners and center of a frame, similar to [4]. The size of each ROI is based on the frame size. Figure 3 depicts the resulting mask for a 4:3 aspect ratio.

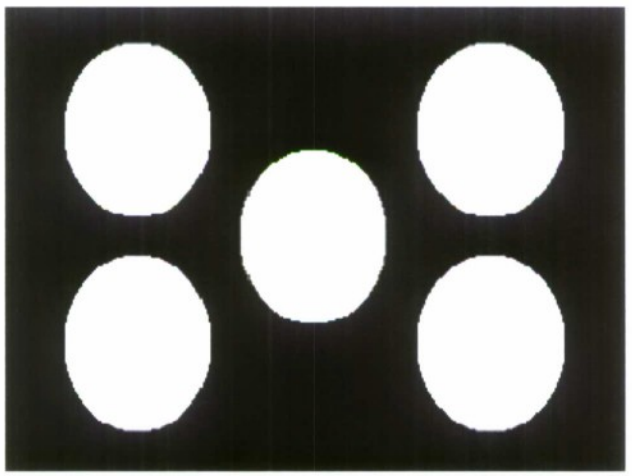


Figure 3: Five ROI for a 4:3 Video Frame

Calculate Bhattacharyya Coefficients

With five regions of interest selected for the current video frame as well as three models, a total of 15 Bhattacharyya coefficients must be calculated. First, each ROI is converted to its approximated pdf (see *Generic Face Models* for methodology). Noting

$$\text{Bhattacharyya Coefficient} = \hat{\rho}(y) = \sum_{h=1}^{m} \sum_{s=1}^{n} \sqrt{\hat{p}_{h,s}(y)\hat{q}_{h,s}}$$

$$\text{Bhattacharyya Distance} = \sqrt{1 - \hat{\rho}(y)}$$

where,

h = histogram bin number for hue
s = histogram bin number for hue

$\hat{p}_{h,s}(y)$ = probability of target candidate bin, [h, s], at frame location y

$\hat{q}_{h,s}$ = probability of target model bin, [h, s]

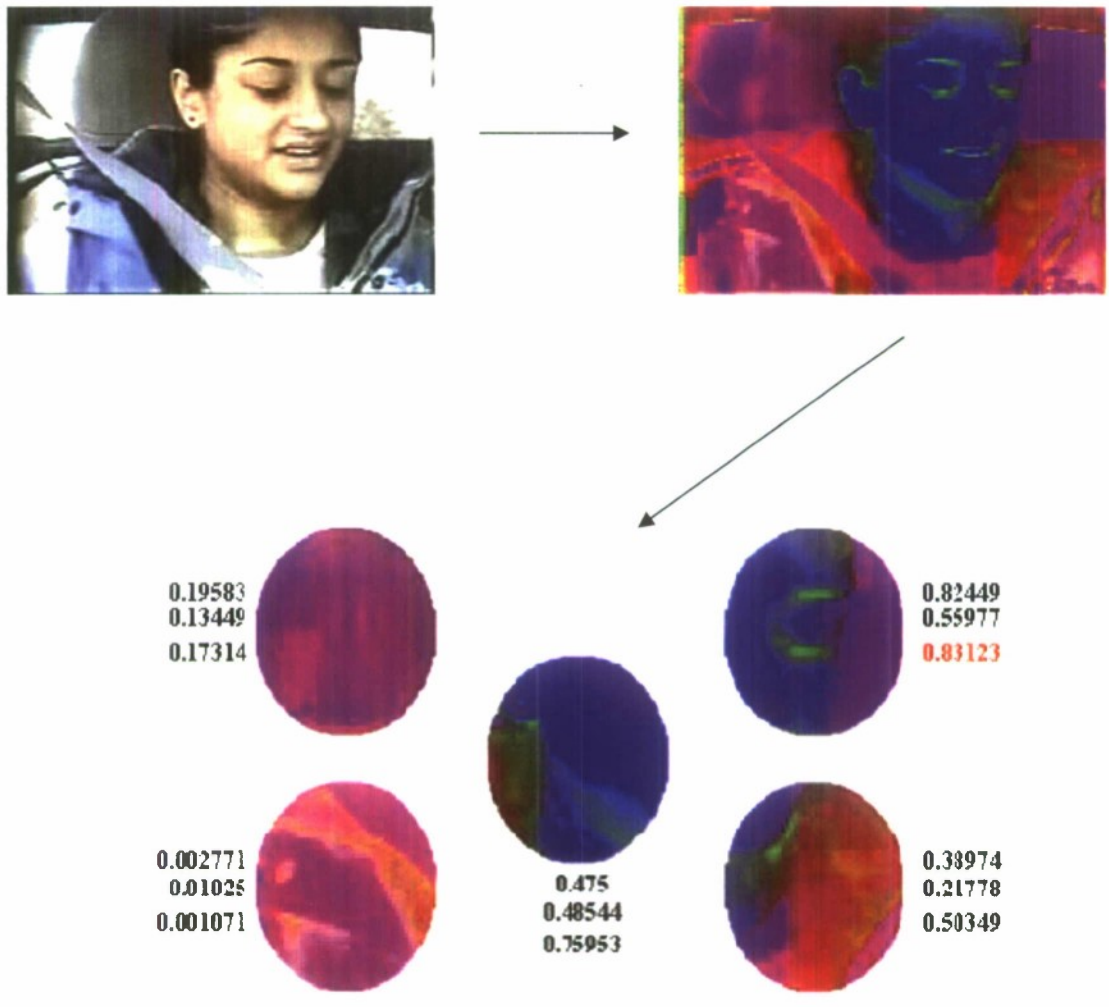


Figure 4: Graphic Representation of Bhattacharyya Coefficient Calculations

The Bhattacharyya coefficient is a measure of two distribution's similarity, where 1 equates to two identieal distributions and 0 equates to two dissimilar distributions. The Bhattaeharyya distanee is just the opposite; a distanee of 0 equates to two identical distributions and a distance of 1 equates to two dissimilar distributions. Values between 0 and 1 represent an equivalent degree of similarity. Therefore, the maximum Bhattaeharyya coeffieient (minimum distanee) of the 15 ealculated is more likely a faee than the other 14. Experimental results suggest a Bhattaeharyya eoeffieient of 0.6 or greater eontains a face within the given ROI. Therefore, a threshold of 0.6 ('face_thresh' of Figure 1) is implemented to aecept/rejeet ROIs. Finally, as depieted in Figure 1, if the maximum Bhattaeharyya eoeffieient is greater than or equal to the threshold, the ROI's parameters are returned to its invocating funetion. A graphieal representation of this module's funetionality ean be seen in Figure 4, where the numbers loeated near the ROIs are their respeetive Bhattaeharyya eoeffieients. For clarity, it should be noted that the

pdfs of the ROIs are not depicted, but the sHSI ROIs depicted near the bottom of Figure 4 are converted to their pdf approximations to calculate the Bhattacharyya coefficients. The largest Bhattacharyya coefficient is 0.83123 and was calculated from the upper-right ROI and the third face model (center face of Figure 2). Note that the individual in Figure 4 is not one of the individuals used for the generic face models.

2.2.2 Extract Face Coordinated

As can be seen in Figure 4, while a calculated Bhattacharyya coefficient can determine if a face is present within a given ROI, it does not guarantee the complete face is within the ROI. Therefore, the *Extract Face Coordinates* module is responsible for expanding the ROI found in the *Find Face* module to that of the complete face. This is accomplished using statistically determined thresholds and simple binary morphology. Figure 5 depicts the top-level design of the *Extract Face Coordinates* module; a more in depth explanation of the module's functionality and results follow.

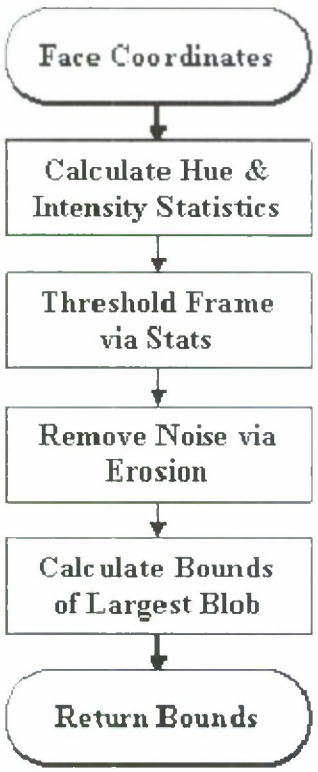


Figure 5: Top-Level Design of Expand Face Coordinates Module

While saturation is invaluable in locating features in a region, it is a poor metric for locating a feature rich region such as a face since saturation values can vary drastically. However, with the exception of shadows, intensity values remain relatively constant for a given region. Additionally, using sHSI, skin hue falls primarily between the values of 0.0 and 0.4. Therefore, both hue and intensity are chosen as metrics for video frame feature discrimination.

It is assumed that a face detected within a given frame is illuminated from a source in front of the facial plane (i.e. not back-lit) and, therefore, a minimal value of 0.2 in intensity is used as a threshold to remove dark/shadowed regions of the passed ROI. The median and minimum value of the ROI's resulting intensity space is then calculated. Similar computations are completed for the shifted-hue space. These statistical values are then used to produce a binarized frame. To remove noise and erroneous connected components, the binarized frame is then eroded. Finally, the largest blob remaining in the frame, and nearest the *Find Face* module's selected ROI, is the face. Therefore, summation about the rows provides the coordinates for the left and right of the face. Using these coordinates, as well as a face height/width ratio of 1.2, a bounding box is created and 'slid' down the rows to find maximum pixel containment. This biases the coordinates for the top and bottom of the face towards the top of the image, removing potential coordinate selection errors based on neck visibility. Specifically, while the typical face has a height/width ratio of approximately 1.2, since this module selects skin, the neck of an individual would affect the extracted bounds. Examples of this processing can be seen in Figure 6.

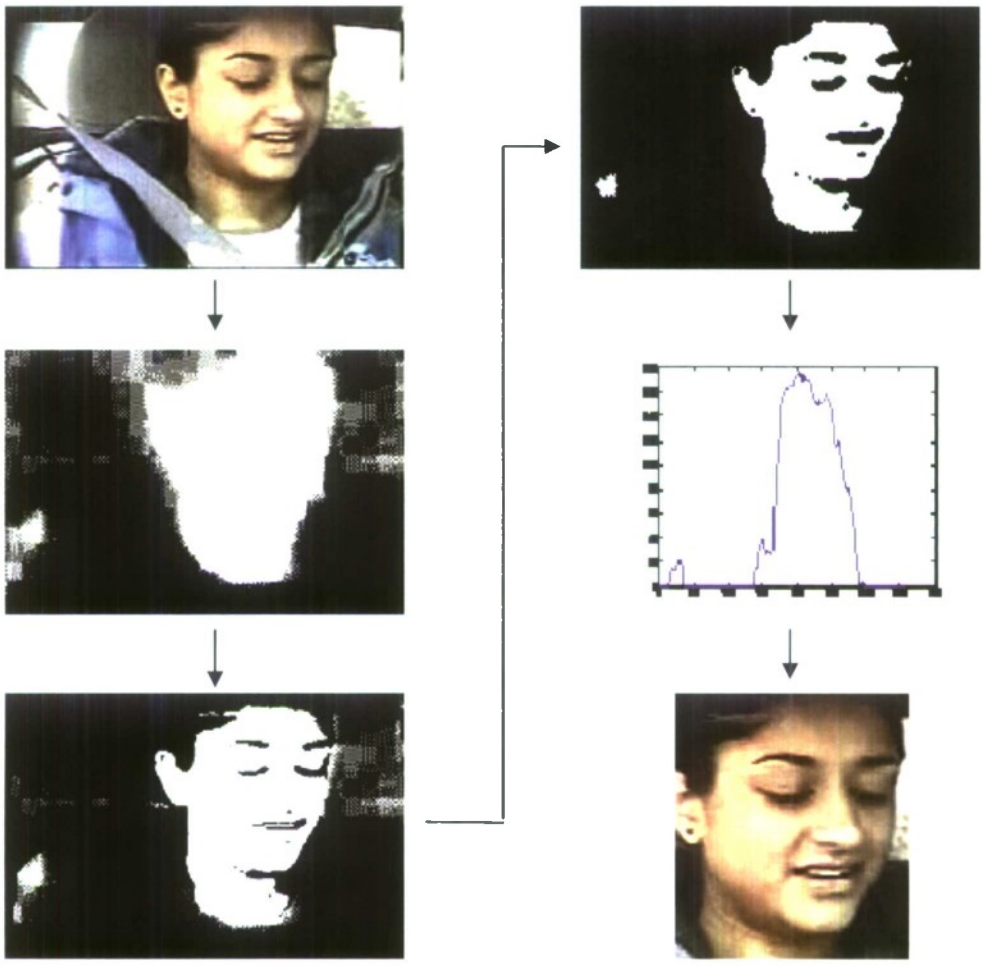


Figure 6: Example of Extract Face Coordinate Processing

Finally, the extracted bounds are then oversized to counteract erosion effects as well as the addition of some background pixels. The addition of background pixels is necessary to provides a less stringent (tight) tracking of the face downstream as well as more robust edge detection, via spatial gradient, for facial scaling (face closer/further).

2.2.3 Create Face Target Model

Using a modified version of the Epanechnikov kernel as a weighting function for the distance from the center of the ROI to a current pixel location, an approximate pdf (histogram) is generated – this is the target model. This module returns the modified pdf, $\hat{q}_{h,s}$, which is an m x n matrix (m shifted-hue bins, n saturation bins). Figure 7 depicts the top-level design of the *Create Face Target Model* module; a more in depth explanation of the module's functionality and results follow.

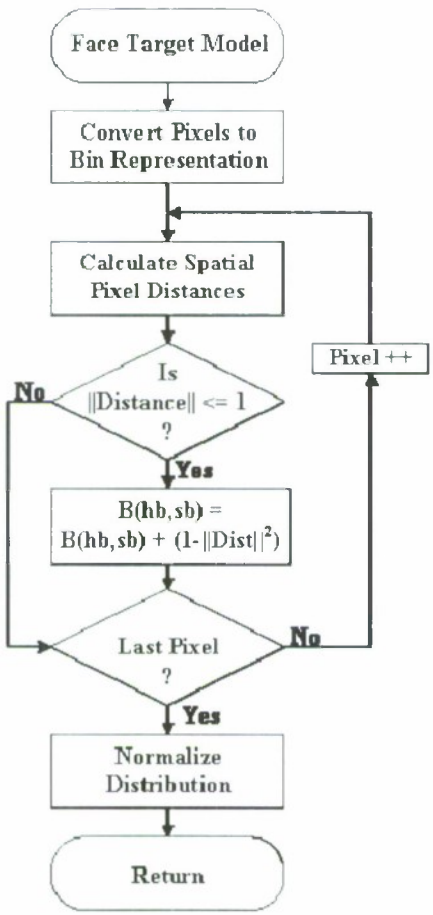


Figure 7: Top-Level Design of Face Target Model Module

To convert the pixels within the ROI to their bin representation, the following formula is used for both the shifted-hue and saturation space:

$$bin(pixel\,value) = floor\left(\frac{pixel\,value \cdot total\,bins}{256}\right) + 1$$

where,

pixel value = unsigned 8-bit integer
total bins = total bins for shifted-hue (or saturation)

The pixel distance is calculated relative to the ROI's center and is always normalized. In other words, the distance from the center of the ROI to the perimeter of the ellipse is equal to 1. This is a crucial step that provides model/candidate scale and rotational invariance. The following formula is used to calculate the pixel distance:

$$D_p(r,c) = \left(\frac{[r,c] - roictr}{roictr}\right)$$

The Epanechnikov kernel requires the squared Euclidean norm, $\|D_p(r,c)\|^2$. However, since the distance is normalized, the square of the norm will always result in a value less than 1 for distances less than 1. Therefore, the pixel distance calculation can be simplified by removing the unnecessary steps, replacing the Euclidean distance with that of the squared norm. Thus,

$$\hat{D}_p(r,c) = \|D_p(r,c)\|^2 = D_p(r,c) \cdot D_p^T(r,c)$$

Finally, all leading constants to the Epanechnikov kernel have been dropped since the output of this module is an approximation of the distribution's pdf (i.e. sums to 1). In other words, the kernel's coefficients would drop out once the distribution was normalized.

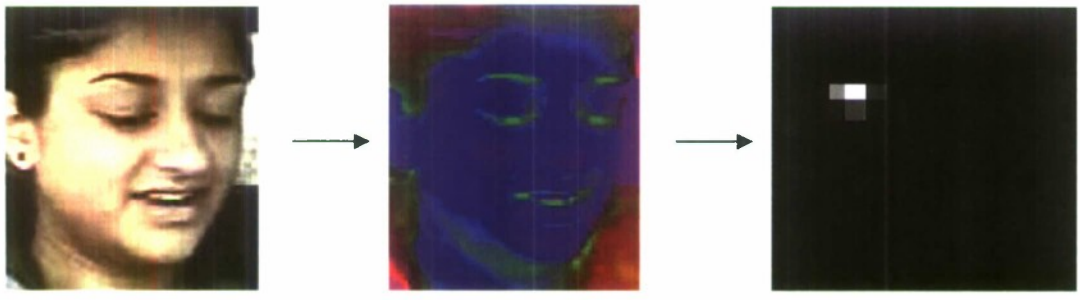


Figure 8: Example of Resulting Face Target Model

The pdf for a given ROI is calculated by adding to a pixel's corresponding bin the weight, $1 - \hat{D}_p$, if the pixel's normalized distance is less than 1 from the center of the ROI. After all pixels are processed, the resulting bins are normalized to generate an approximation to the distribution's true pdf. Figure 8 shows an example of the module's resulting approximation pdf (face target model).

Note: for elarity, the pdf was scaled by 1/max[pdf] as the bin values are too low to see otherwise.

2.2.4 Extract Lip Coordinates

This module is similar to the *Extract Face Coordinates* module but uses saturation and intensity rather than hue and saturation. Intensity is chosen as a metric due to the low values at the mouth's opening, regardless of it being open or elosed, as well as most lighting eonditions. Speeifieally, other than a light source being plaeed in the mouth or on a point parallel to the mouth's opening plane, the lip opening will always have lower intensity values than the surrounding lip region. Additionally, the algorithms used for this module are much more complex to aceount for potential feature noise (nose, shadows, faeial hair, ete.). Figure 9 depicts the top-level design of the *Extract Lip Coordinates* module; a more in depth explanation of the module's functionality and results follow.

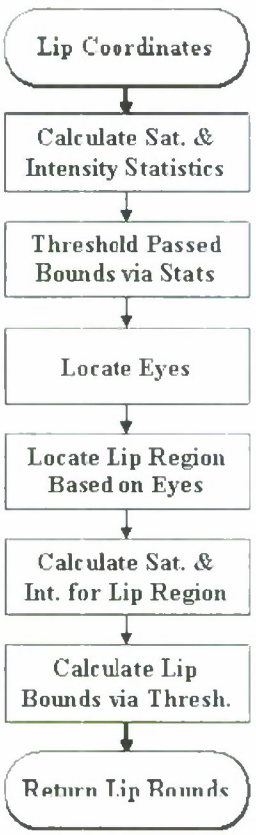


Figure 9: Top-Level Design of Extract Lip Coordinates Module

The ROI reeeived by this module contains only a face with a perimeter consisting of baekground noise. Therefore, seleeting a sub-region, eentered about the ROI, assures statistieal analysis will only take into aeeount the faee. Using the statisties of mean and standard deviation for both saturation and intensity, several thresholds can be implemented to remove the skin of the faee, leaving only the faeial features within the frame, whieh is then binarized. Next, the left and right side of the faee ean be determined

by summing the ROI about its rows, resulting in a spatial signal that can be analyzed for its maximums – left and right edge of face. This data is then used to generate a face mask, which has a predefined shape but its height and width is determined by the width of the face and the facial height/width ratio of 1.2.

However, noise is still prevalent at this point, requiring larger features to be located – the eyes. This is accomplished by first dividing the ROI into a top and bottom half and then locating the two maximum valued distributions from the top binarized frame. After the eyes are located the region (bridge of the nose) can be calculated. This allows separate calculations of the left and right eye to determine the facial tilt angle. The tilt angle, combined with knowledge of the eye height and facial width, allows calculation of the expected location of the lips. Additionally, this knowledge allows additional noise removal by determining areas of no interest, which can then be set to zero, preventing contamination to the lip data. Figure 10 provides select examples of this processing.

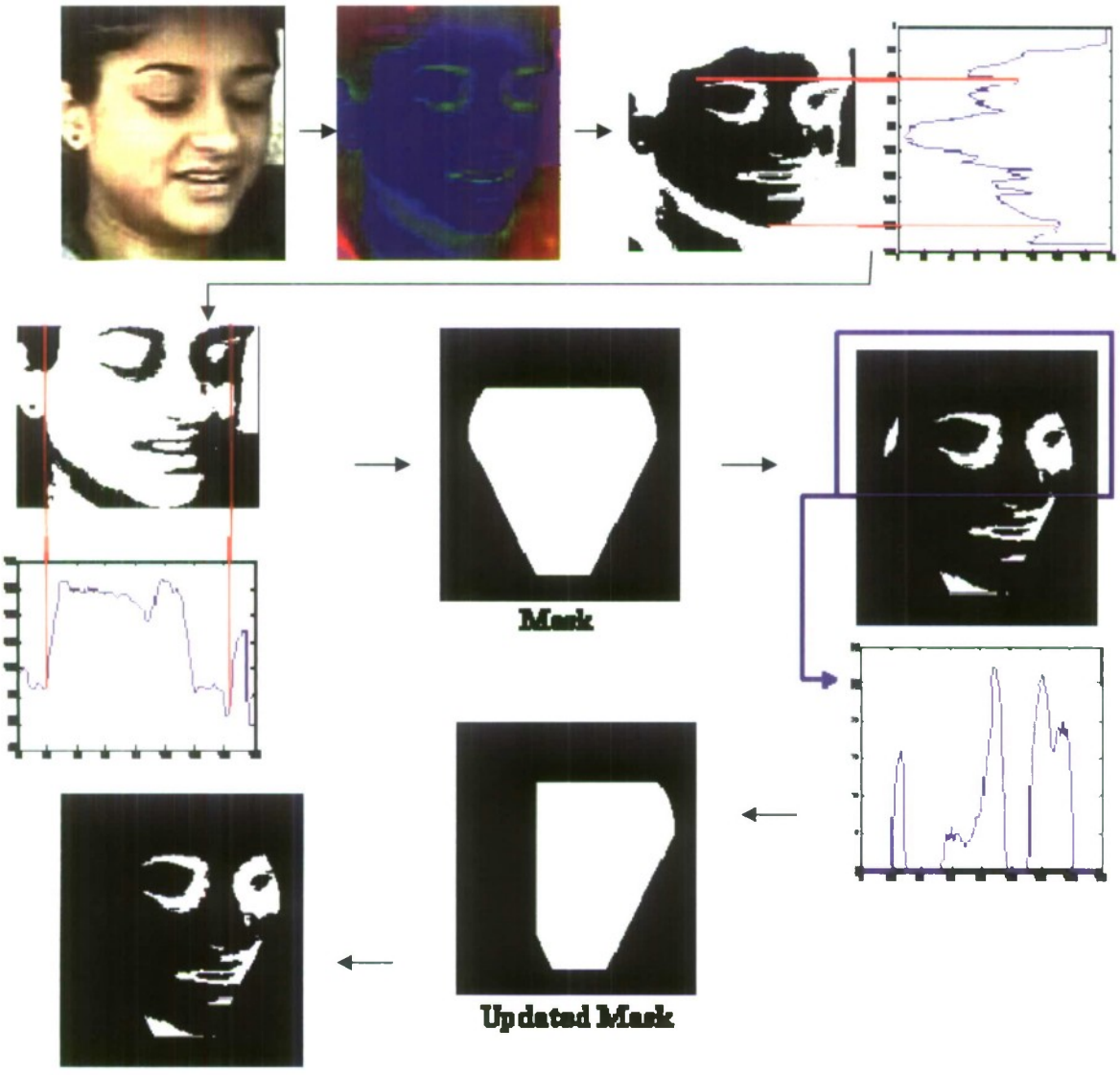


Figure 10: 1st Stage of Lip Coordinates Module

With the expected location of the lips ealculated and additional noise removed, analysis of the lower face ean begin. Summing about the eolumns of the lower face, maximums are located within the bounds of the expeeted loeation of the lips. Onee found, the proeess is repeated about the rows of the lower faee. These two processes provide the initial lip bounds.

However, noise eould still be present whieh could potentially increase the lip bounds. This noise includes facial hair, dimples, the shadow of the nose, etc. Therefore, the mean and standard deviation for both saturation and intensity are ealeulated for the eurrent lip region. Using statistieal thresholds, the eurrent lip region is reanalyzed to remove additional noise and tighten the bounds to just the lip opening. Next, the new bounds are compared to the known faeial tilt and expeeted lip loeation, if the lip width is not approximately eentered to the faee or is eonsidered to large/small for the faee, a region growing algorithm is implemented. Finally, the lip opening bounds are extracted and then enlarged on all sides by $1/20^{th}$ the width of the lip bounds. Additionally, the bottom of the lip bounds is set as 2/3rds that of the lip width, ensuring a ROI size that will handle the changes of the mouth opening and elosing, as well as minor variations in facial tilt. Figure 11 displays these final proeessing steps as well as the ROI returned from the module.

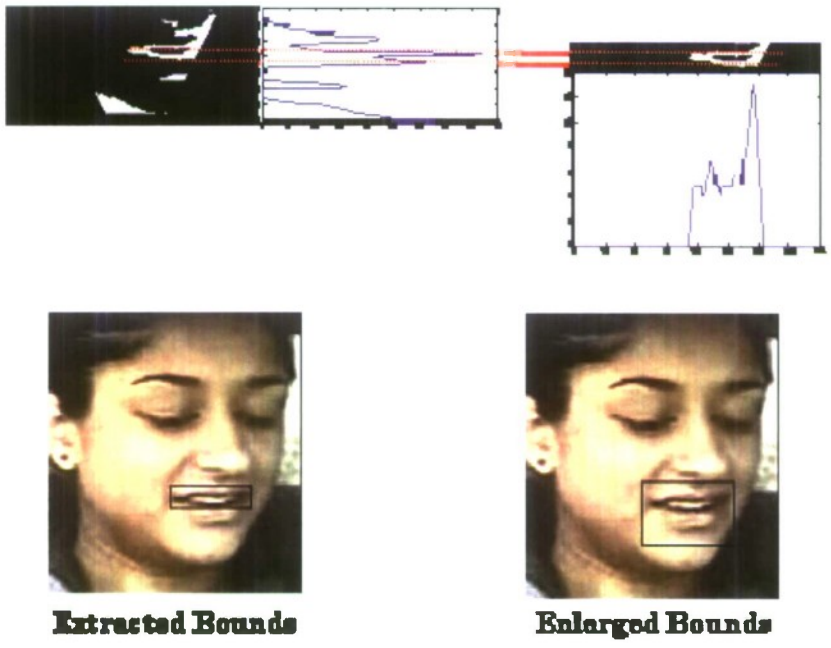


Figure 11: Final Stage of Lip Coordinates Module

2.2.5 Create Lip Target Model

This module is identical to the *Create Face Target Model* module. Figure 12 depicts the top-level design of the *Create Lip Target Model* module; examples of the module's results can be seen in Figure 13.

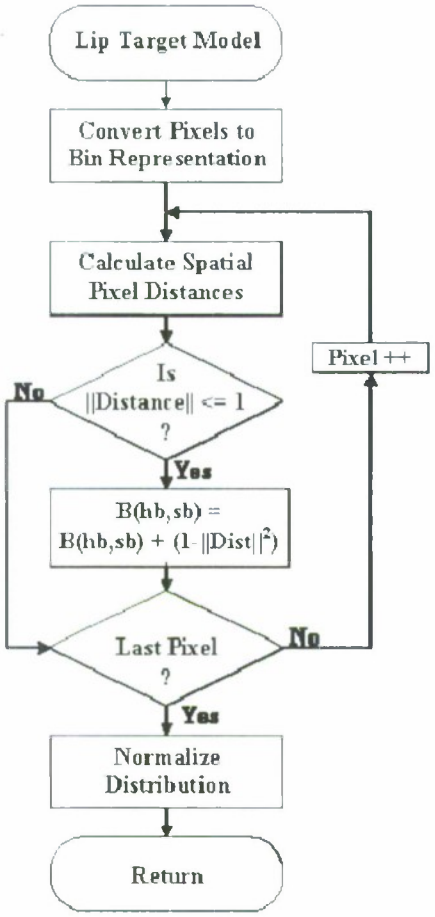


Figure 12: Top-Level Design of the Lip Target Model Module

Figure 13: Examples of Lip Target Model Results

2.2.6 Track Face

This module is built around the mean-shift algorithm and is responsible for not only the tracking of the face but scaling as well. Specifically, this module will adjust the ROI bounds for each frame based on 3-dimensional facial movements (left/right, up/down, in/out). Figure 14 depicts the top-level design of the *Track Face* module; a more in depth explanation of the module's functionality and results follow.

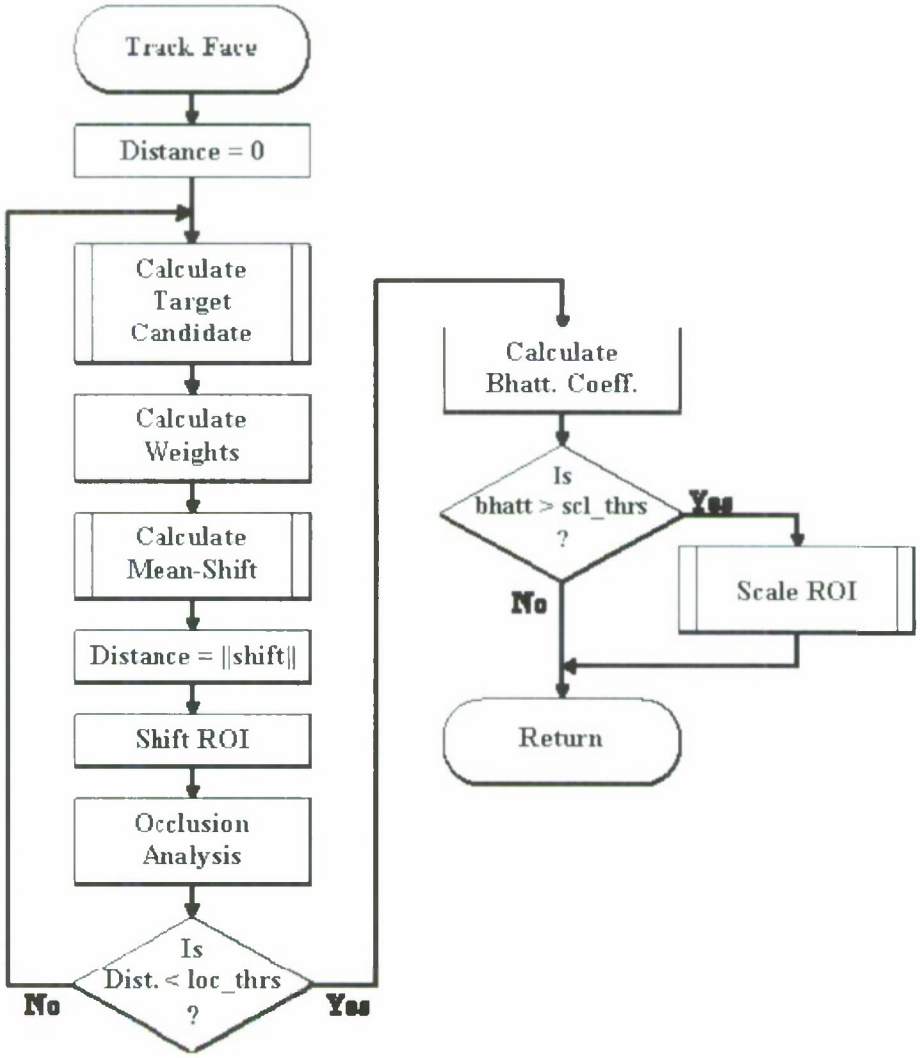


Figure 14: Top-Level Design of the Track Face Module

Calculate Target Candidate function

This module is identical to the *Calculate Target Model* module but returns $p_{h,s}$, the approximated pdf of the candidate ROI.

Calculate Weights

The mean-shift weights are determined via the formula

$$w_{h,s} = \sum_{h=1}^{m}\sum_{s=1}^{n}\sqrt{\frac{q_{h,s}}{p_{h,s}}},$$

and is used, in eonjunetion with the systems eurrent location, to ealculate the direction of shift (traeking). An example of weight ealeulation can be seen below, based on histograms eomprised of 2 bins for shifted-hue and 2 bins for saturation. Note that if the model's probability for a partieular bin eombination is larger than that of the candidates, a large weight will result.

0.1	0.5
0.2	0.2

$q_{h,s}$

0.2	0.3
0.1	0.4

$p_{h,s}$

$$\text{weight} = \begin{array}{|c|c|}\hline \sqrt{0.5} & \sqrt{1.3} \\ \hline \sqrt{2} & \sqrt{0.5} \\ \hline \end{array}$$

$w_{h,s}$

Calculate Mean-Shift and Shift ROI

The new location is determined using the weights derived above and the pixel locations of the target eandidate, $x_{pd,h,s}$, normalized about the ROI's eenter as follows:

$$\text{Loeation shift} = \frac{\sum x_{pd,h,s} w_{h,s}}{\sum w_{h,s}}$$

Specifically, $x_{pd,h,s}$ is converted similarly to that of *Face Target Model* module except that pixel direetion is required and, therefore, pixel distanee must be ealculated. Pixel distanee is defined for this module as

$$D_p(r,c) = \left(\frac{[r,c] - roictr}{roictr}\right).$$

The conversion of pixel values to bins is identical to that of the *Face Target Model* module. Finally, if the norm of the pixel distanee is less than or equal to 1, the normalized pixel direetion is multiplied by the weight eontained in the eorresponding bins of the weight matrix. Thus, the target is traeked by attempting to match the pdf's of the model and eandidate distributions via direetion, based on pixel loeation, and the pixel's shifted-hue and saturation, via its eorresponding bins.

Finally, once the loeation shift is ealculated, the result must be multiplied by ½ the size of the target candidates loeation ROI size. This removes the normalization allowing the correet shifting of the ROI loeation. In other words, if the result was not 'denormalized', the shift would remain eentered about the same pixel as the norm of the distance would be less than 1. The following subseetion provides a sample ealeulation of the mean-shift as well as the location shift to promote a better understanding of the proeesses involved.

Finally, several tracking examples can be seen in Figure 15 and Figure 16. Additionally, a visual example of these calculations can be found on page 21.

Scale ROI Sub-Module

Figure 17 provides a schematic representation of the top-level design of the *Scale ROI* sub-module. The target ROI is oversized and occlusion analysis is calculated to ensure the oversized ROI is within the frame's bounds. Three scaling factors are then calculated, representing an oversized, undersized, and unchanged ROI size. Next, using a ring around the ROI's perimeter, the average optical flow magnitude, based on the intensity gradient and normalized about the ROI's center, is calculated. Finally, the minimum difference between the three scaled optical flows and that of the model's is considered the most similar scale. The ROI bounding box corners are updated and the process returns if either the maximum scale iterations has been reached or the resulting scale = 'same', otherwise it repeats.

Occlusion Analysis function

This sub-module is responsible for ensuring the new ROI location is within the frames limits. Specifically, if any part of the new ROI is outside the dimensions of the image, this function truncates the offending portion.

Calculate Bhattacharyya Coefficient

The Bhattacharyya coefficient is used to determine when and if the sub-module, *Scale ROI*, should be invoked. If the calculated Bhattacharyya coefficient remains above 0.6, but falls below 0.8, *Scale ROI* is invoked.

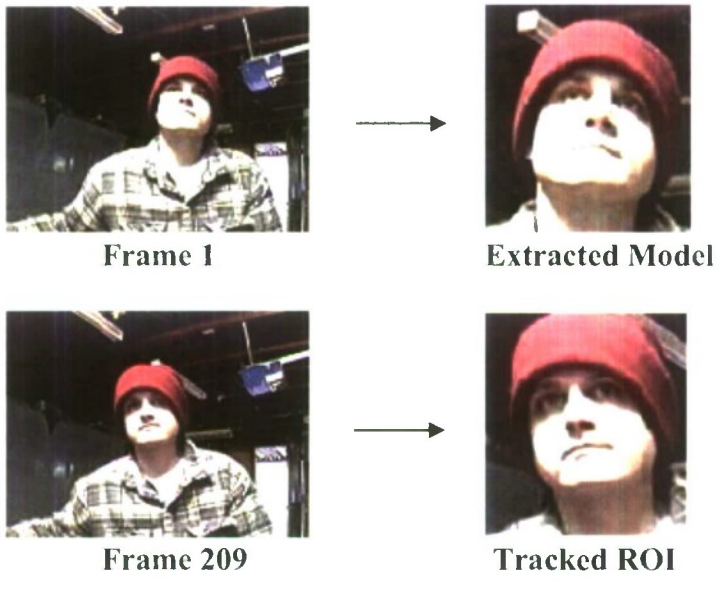


Figure 15: Example 1 of Face Tracking

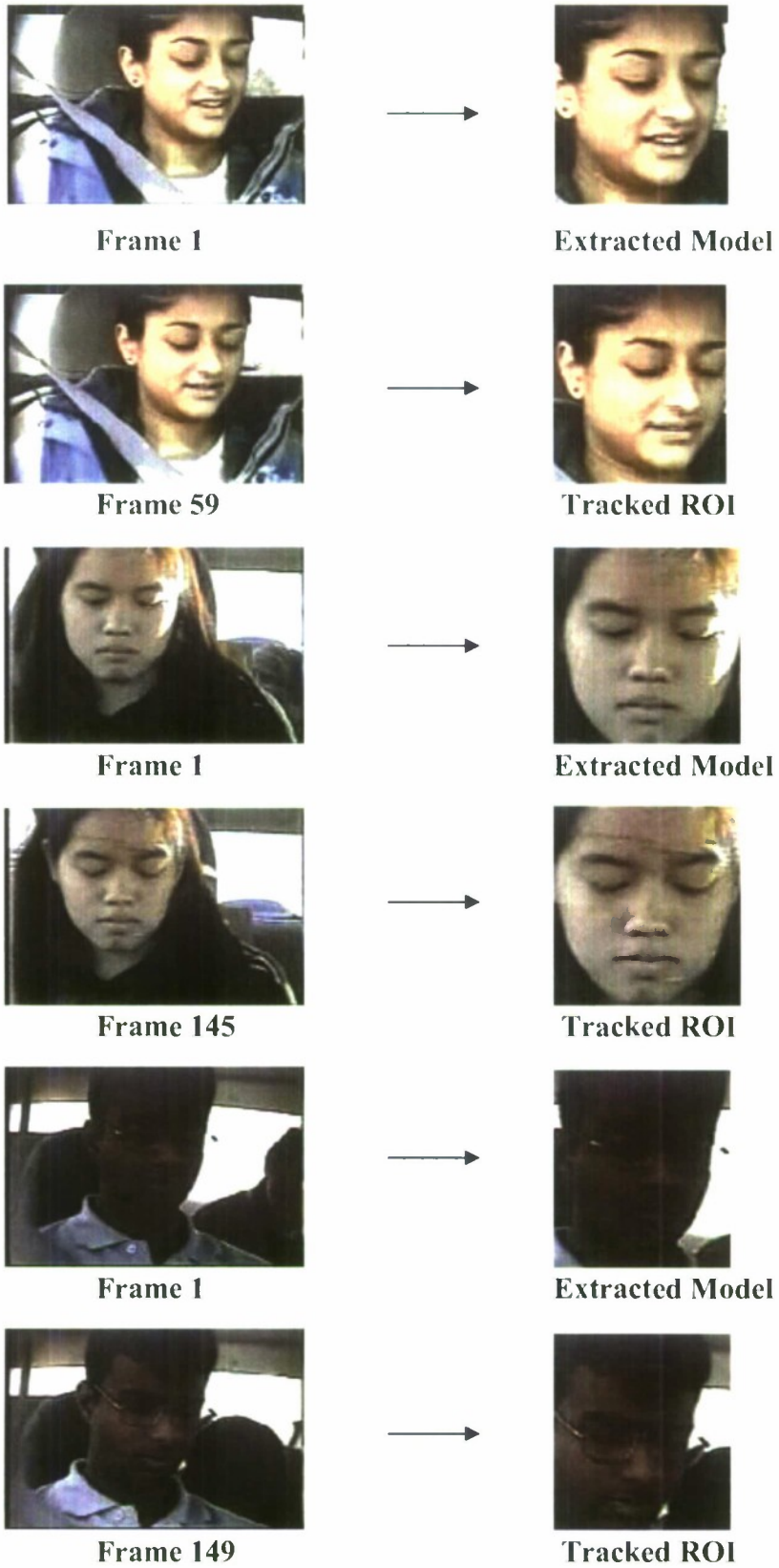


Figure 16: Example 2 of Face Tracking

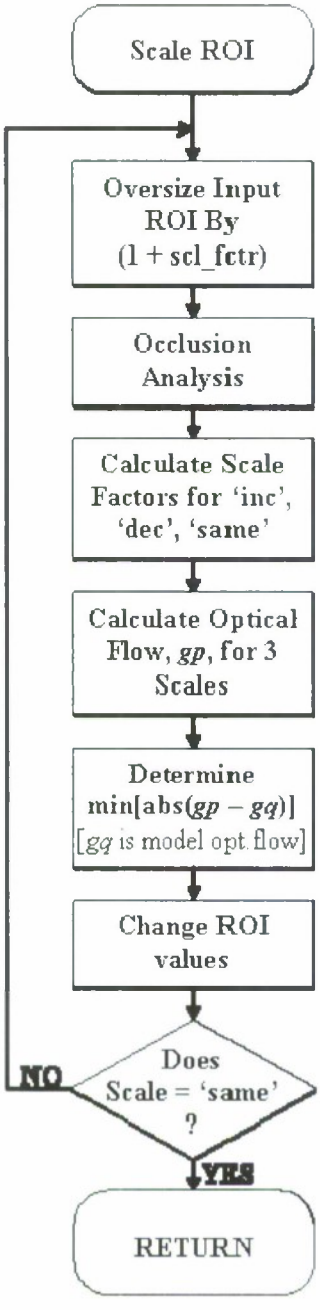


Figure 17: Top-Level Design of the Scale ROI Sub-Module

A Visual Calculation of the Mean-Shift and Location Shift

The following calculations use frame 1 and frame 5 of the video file "EF3_35D_bl.avi." A three frame delay is utilized to increase the differences between their distributions. Frame 1 is the target model, $\hat{q}_{h,s}$, while frame 5 is the target candidate, $\hat{p}_{h,s}$.

To begin, the approximated pdf's for both the target model, $\hat{q}_{h,s}$, and candidate, $\hat{p}_{h,s}$, must be calculated, which was discussed in detail on pages 9, 10 and 16. Next, the weight matrix must be calculated using the following equation,

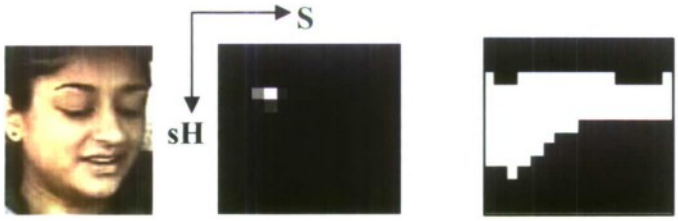
$$w_{h,s} = \sum_{h=1}^{m}\sum_{s=1}^{n}\sqrt{\frac{q_{h,s}}{p_{h,s}}}$$

Figure 18 displays the pdf's and weight matrices for the model as well as the candidate for the 3 required iterations of the mean-shift algorithm. Note that for clarity, the pdf's and weighting matrices have been scaled. Additionally, e-weight refers to the effective weight matrix – a visual manipulation of the weight matrix to display all bins with weights greater than zero (equivalent explanation for e-pdf).

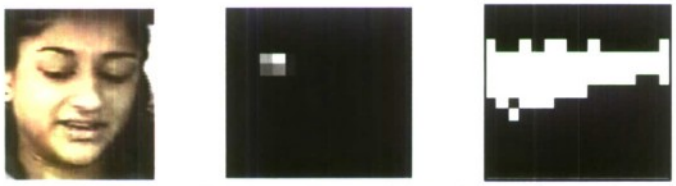
Once the weight matrix has been calculated, the location shift is calculated using

$$\text{Location shift} = \frac{\sum x_{pdf,h,s} w_{h,s}(x)}{\sum w_{h,s}(x)},$$

this can be further understood by referencing Figure 19. The left image is the current ROI's shifted-hue (red) and saturation (green) bin image – the number of bins (discrete ranges) for both shifted-hue and saturation is 16, therefore the image is comprised of 16 diserete pixel levels for each. The right illustration represents the normalized pixel distance for the ROI – note that the outer edge has a norm of 1, hence normalized. Therefore, for each pixel, one can calculate the distance from the center of the normalized ROI, as well as access the bin values stored within that partieular pixel. Thus, to calculate the location shift, the distanee is multiplied by the weight for a given bin combination (stored within the current pixel) and then, onee all pixels within the candidate ROI have been calculated, the result is scaled by the total value of all weights used. In essence, the pixel's values (bin numbers) are inputs to a look-up-table (LUT), the weight matrix, whieh ean be considered a significanee metric, the larger the weight the more signifieant the pixel value is to the model ROI, and therefore the normalized direetion is more heavily biased.

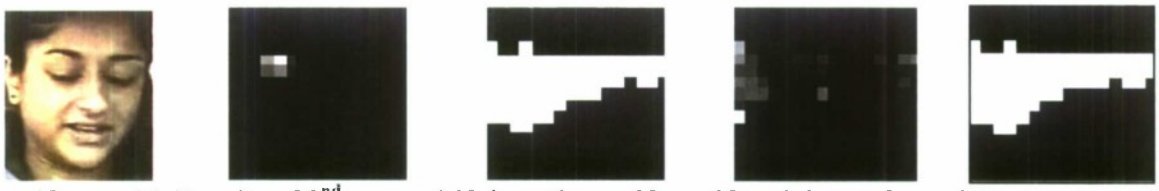


Frame 1: Determined face ROI, pdf, and e-pdf matrices



Frame 15: Starting location, pdf, and e-pdf matrices

Frame 15: Results of 1st mean-shift iteration, pdf, e-pdf, weight, and e-weight matrices



Frame 15: Results of 2nd mean-shift iteration, pdf, e-pdf, weight, and e-weight matrices

Frame 15: Results of the 3rd mean-shift iteration, pdf, e-pdf, weight, and e-weight matrices

Figure 18: PDF and Weight Matrices for the Visual Example

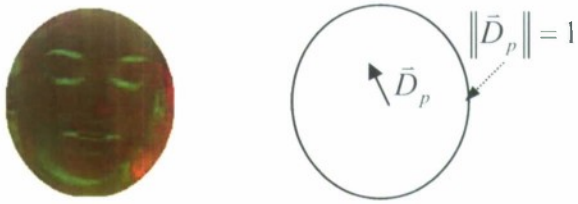


Figure 19: Pixel Distance and Bin Values for the Visual Example

2.2.7 Track Lips

This module is identical to the *Track Face* module. Figure 20 depicts the top-level design of the *Track Lips* module; examples of the module's results can be seen in Figure 21.

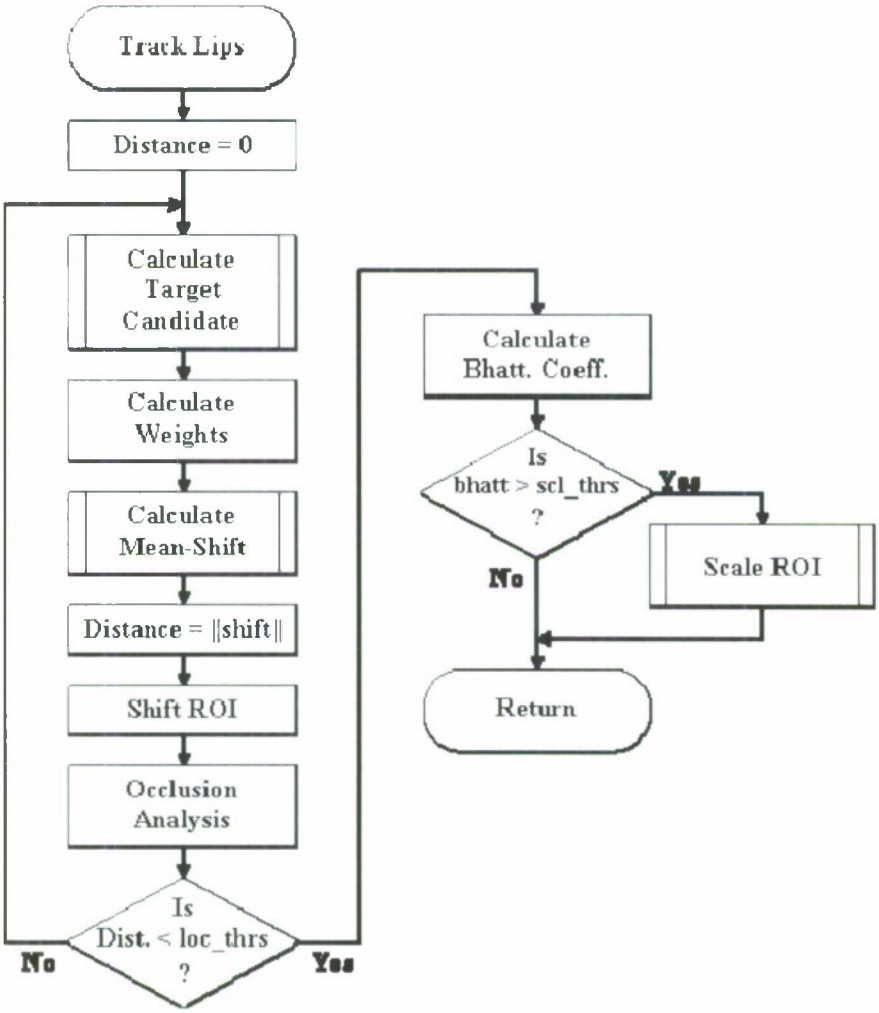
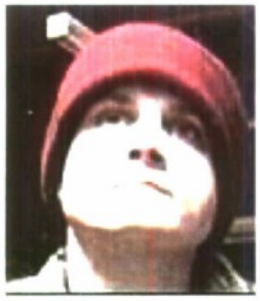
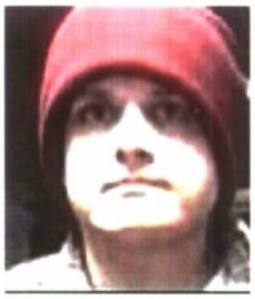


Figure 20: Top-Level Design of the Track Lips Module


Frame 1


Extracted ROI


Frame 409


Tracked ROI


Frame 1


Extracted ROI


Frame 93


Tracked ROI


Frame 1


Extracted ROI


Frame 43


Tracked ROI


Frame 1


Extracted ROI


Frame 145


Tracked ROI

Figure 21: Examples of Lip Tracking

3. Lip Reader System Examples

Example 1

Figure 22 displays three frames from the traeking results for the faee and lips of the "my_vid15.avi" video. The video is 14 seeonds in duration and is reeorded at 30 fps. The runtime for the *Lip Reader System* ean only be approximated at this time, as all modules are not currently series eonneeted. Speeifieally, the system is separated into faee and lip modules; where the faee is extraeted from the video souree and then outputted to an .avi file, and the lip module then analyzes this file to extraet the lips. Therefore, while the times ean be given for the two separate proeesses, the finalized system will be faster by removing the storage aspeet of the intermediate step. Using a desktop eomputer with 1GB of memory and a 2.93 GHz Intel Celeron CPU, the following times were reported for the two modules using Matlab R2006a:

Table 1: Runtime for "my_vid15.avi" (duration = 14 seconds)

Face Location, Tracking and File Writing	**60.7 seconds**
Lip Location and Tracking	**7.77 seconds**

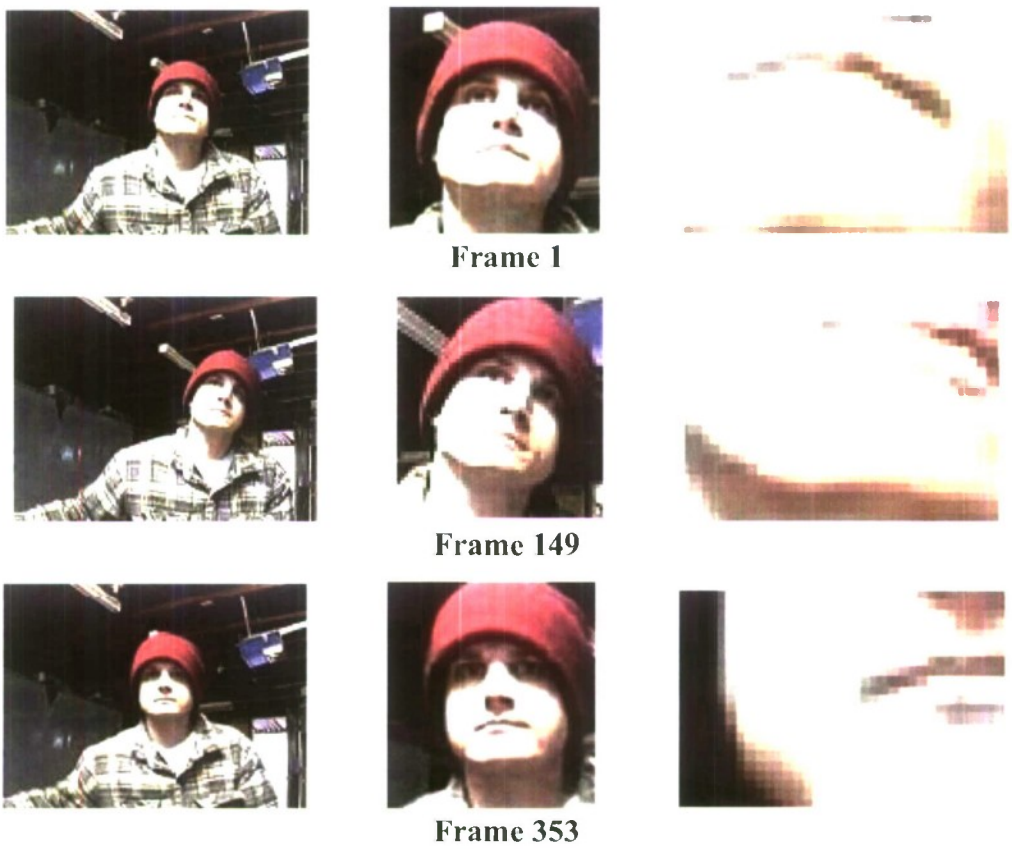


Figure 22: Tracking of Face and Lips for "my_vid15.avi"

While the runtime in Matlab is approximately 68 seconds, or 20% that of real-time, porting this system to C will provide a minimum of 10 times faster processing. In other words, once the system is finalized and ported to C, the system can run at a minimum of 2 times faster than real-time.

Referencing Figure 22, the lip tracker performed poorly in frames 149 and 353 due to quick movements of the head. This poor performance, however, is acceptable as one should expect the user of a system such as this to remain relatively stationary when speaking. The next example shows such a case.

Example 2

Figure 23 displays three frames of the tracking results for the face and lips of the "EF3_35D_bl.avi" video. The video is 2 seconds in duration and is recorded at 30 fps.

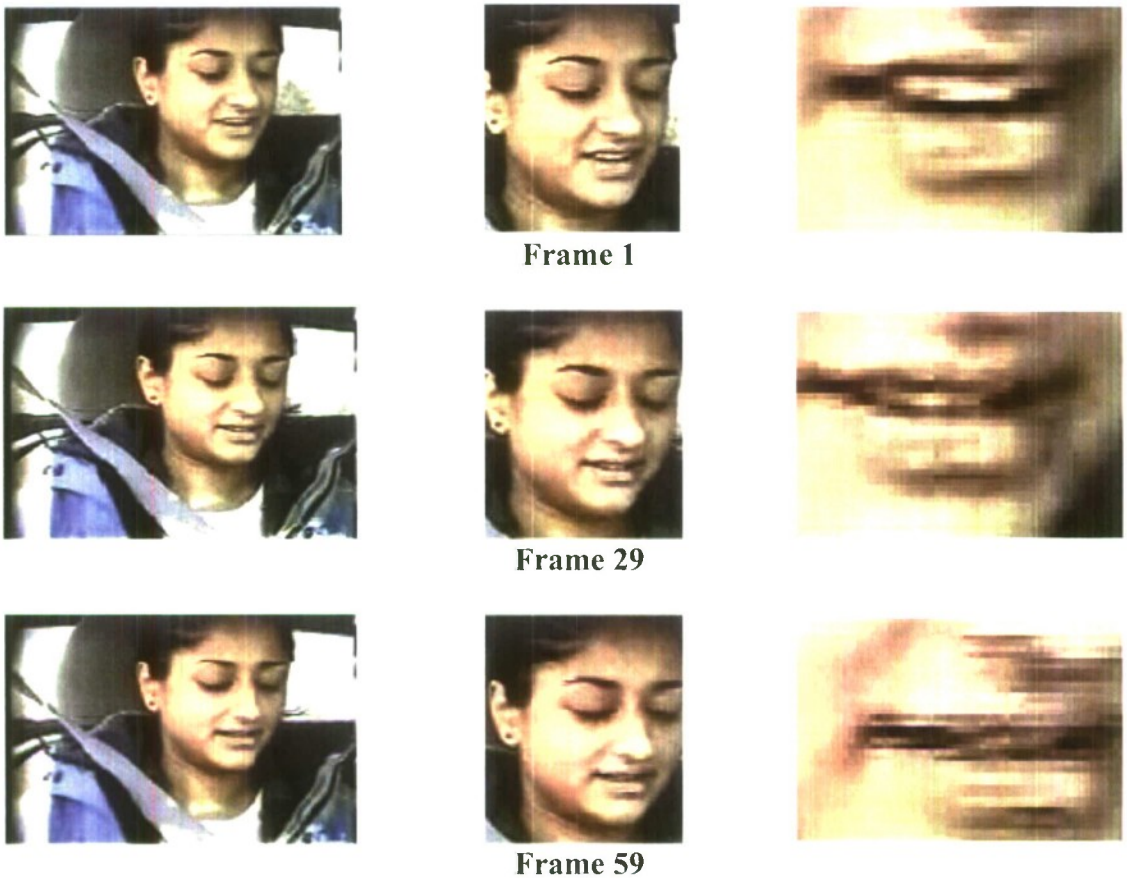


Figure 23: Tracking of Face and Lips for "EF3_35D_bl.avi"

Using a desktop computer with 1GB of memory and a 2.93 GHz Intel Celeron CPU, the following times were reported for the two modules using Matlab R2006a:

Table 2: Runtime for "EF3_35D_bl.avi" (duration = 2 seconds)

Face Location, Tracking and File Writing	**16.2 seconds**
Lip Location and Tracking	**2.56 seconds**

While the runtime in Matlab is approximately 19 seconds, or 11% that of real-time, porting this system to C will provide a minimum of 10 times faster processing; thus, the system can run at a minimum of real-time.

The increased processing time (fps processed) for the "EF3_35D_bl.avi" video versus that of the "my_vid15.avi" video is due to both an inereased ROI size as well as edge occlusion. Since the mean-shift algorithm is an iterative process, anytime a target candidate is occluded by a frame edge (or object passing in front of target), the mean-shift iteration count will increase – attempting to locate the target (maximize the Bhattacharyya coeffieient). A maximum of ten iterations of the mean-shift algorithm is allowed per frame, but this limit is only approached for occluded targets. On average, a video sequence requires only 2.1 iterations per frame. However, this average can readily double (or more) if the target is frequently occluded.

4. Future work

Built upon the previous results, next we extract lip parameters that will be input to a pattern classification engine to perform automatie speech reeognition,

"This work was sponsored by the Department of the Navy, Office of Naval Research, under Award # N00014-06-1-1111."

References

[1] Lee, B., Hasegawa-Johnson, M., Goudeseune, C., Kamdar, S., Borys, S., Liu, M., Huang, T.: AVICAR: Audio-Visual Speech Corpus in a Car Environment. In: INTERSPEECH2004-ICSLP, 2004.

[2] Zhang, X., Montoya, H.A., Crow, B.: Finding Lips in Unconstrained Imagery for Improved Automatie Speech Reeognition. In: Proe. 9th International Conference on Visual Information Systems, VISUAL 2007, pp. 185-192, 2007.

[3] Comaniciu, D., Ramesh, V., Meer, P.: Kernel-Based Object Tracking. In: IEEE Transaetions on Pattern Analysis and Machine Intelligence, Volume 25, Issue 5, Page(s):564 – 577, May 2003.

[4] Comaniciu, D., Ramesh, V.: Robust Deteetion and Traeking of Human Faces with an Active Camera. In: Proe. Third IEEE International Workshop on Visual Surveillance, Page(s):11 – 18, July 1 2000.

HONORS PROJECT

Drop Test Reliability of Lead-Free Solder Interconnections

Project Investigator:

Jianbiao Pan
Industrial and Manufacturing Engineering
California Polytechnic State University
San Luis Obispo, CA

Abstract

This paper presents the drop test reliability of 0.5mm pitch lead-free chip scale packages (CSPs). Fifteen 0.5mm pitch CSPs were assembled on a standard JEDEC drop reliability test board with Sn3.0Ag0.5Cu lead-free solder. Eight boards were edge-bonded with a UV-cured acrylic; eight boards were edge-bonded with a thermal-cured epoxy; and twelve boards were assembled without edge bonding. Half of the edge-bonded test boards were subjected to drop tests at a peak acceleration of 1500G with a pulse duration of 0.5ms, and the other half subjected to drop tests at a peak acceleration of 2900G with a pulse duration of 0.3ms. Half of the test boards without edge bonding were subjected to drop tests at a peak acceleration of 900G with a pulse duration of 0.7ms, and the other half subjected to drop tests at a peak acceleration of 1500G with a pulse duration of 0.5ms. Two drop test failure detection systems were used in this study to monitor the failure of solder joints: a high-speed resistance measurement system and a post-drop static resistance measurement system. The high-speed resistance measurement system, which has a scan frequency of 50KHz and a 16-bit signal width, is able to detect intermittent failures during the short drop impact duration. Statistics of the number of drops to failure for the 15 component locations on each test board are reported. The effect of component position on drop test reliability is discussed. The test results show that the drop test performance of edge-bonded CSPs is five to eight times better than the CSPs without edge bonding. However, the drop test reliability of edge-bonded CSPs with the thermal-cured epoxy is different from that with edge-bonded CSPs with the UV-cured acrylic. The solder crack location and crack area are characterized with the dye penetrant method. The fracture surfaces are studied using scanning electron microscopy (SEM).

I. Introduction

Mobile and handheld electronics devices such as digital cameras, cell phones, and personal digital assistants (PDAs) are prone to be dropped in their lifetime. The drop event may result in failure of solder joints inside these devices. Recently the European Union (EU) Restriction of Hazardous Substances (RoHS) and other countries' lead-free directives banned the use of lead in consumer electronics products. Thus, it is critical to study the drop test reliability of lead-free solder joints.

There has been a significant amount of research done in the last few years on drop impact reliability. The JEDEC standard JESD22-B111 [1] for the board level and related standards [2, 3] for subassembly level have been developed for drop testing handheld electronics. Lim and Low [4] proposed a method to examine the drop impact responses of portable electronic products at different impact orientations and drop height. The impact behavior has been studied at the product level [5]. After comprehensive drop tests, failure analysis, and simulations on two ball grid array (BGA) packages at the board level, Tee, et al. [6] developed a life prediction model for board level drop testing. The effect of different solder alloy compositions on drop reliability has been studied by Syed, et. al. [7]. Since SnAgCu (SAC) solder alloy performs poorly compared with SnPb solder under drop test, several studies have been done to improve the reliability of lead-free solder joints by adding micro-alloying additions [8, 9] or lowering Ag content [10].

Underfill materials were originally developed to improve the solder joint reliability of ball-grid array (BGA) and flip chip packages during temperature cycling. Recently studies have

shown that underfill can improve drop test reliability as well [11, 12]. However the application of underfills increases both the cost of production and assembly cycle times in manufacturing and this must be considered against the reliability improvements. To reduce the costs of underfill application, corner bonding and edge bonding processes have been developed. In the corner bonding process, the adhesive is applied near the package corners before BGA or CSP packages are placed and reflowed. In edge bonding processes, the adhesive is applied after the BGA or CSP packages are placed and reflowed. The reliability of corner-bonded CSPs has been investigated [13, 14].

Failure detection systems and failure criteria used in the literature vary widely. There are three main failure detection methods used in drop test reliability: post-drop (static) resistance measurement [12, 15], event detection [16], and in-situ high-speed data acquisition [17]. The post-drop resistance measurement method measures resistance of solder joints after each drop. The event detection method determines if a failure event temporarily occurs during a single drop. The in-situ high-speed data acquisition method measures the dynamic resistance of solder joints during and after the drop impact and board vibrations. Different researchers have used different failure criteria, for example, a resistance threshold of 300Ω [18], 1000Ω [1], or 1500Ω [16], a resistance change of 10% [12], or 20% [15]. In a sense, all of these criteria are subjective, because, at this time, no scientific research has been done on the interconnection failure criteria. Determination of appropriate failure criteria is extremely important in order to observe first failures and when failures advance to different failure stages [17]. This variety of failure detection systems and failure criteria used by different researchers make the comparison of results difficult. This study discusses the reliability of CSPs in drop impact, with and without edge-bonded underfill, using two failure detection systems and presents the component failure sequence as observed by each system.

II. Test Vehicle Design and Assembly

The test vehicle was designed according to the JEDEC standard [1]. It uses an eight-layer FR4 material board with a size of 132mm by 77mm and a thickness of 1 mm. The component used was 0.5mm pitch Amkor CSP having 228 I/O and with a size of 12mm by 12mm. The CSP has daisy-chained connections with an input and output trace located at one package corner. The boards have Organic Solderability Preservatives (OSP) surface finish on non-solder mask defined (NSMD) pads, while the components have electro-plated nickel-gold surface finish on solder mask defined (SMD) pads. The test vehicle with components assembled is shown in Figure 1.

Sn3.0Ag0.5Cu (SAC305) Multicore 318LF lead-free solder paste (Type 3) was stencil printed using a DEK machine through a 4 mils thick electro-polished stencil with 12 mils square apertures. Solder paste height and volume were measured by a CyberOptic machine to ensure high printing quality. The component was picked and placed by a Siemens F5 machine. A Heller EXL1800 oven with seven heating zones and one cooling zone was used for solder reflow. The reflow oven processing was done in air. The reflow profile is shown in Figure 2.

Post-assembly cross-sectioning and SEM showed good solder joints with some small voids as shown in Figure 3. Visual and X-ray inspection showed shiny, round and well collapsed solder joints with no bridging.

The test boards were divided into three cells, one of which was edge-bonded with a thermal-cured epoxy, one of which was edge-bonded with a UV-cured acrylic, and the third cell

having no edge-bonding. The edge-bond was applied on all four package corners by an Asymtek Century series machine. The edge bond had an average length of 3.81 mm (150 mils) along each side (per corner) after the assembly process was complete, and had an average length of 1.2 mm fillet leg after being cured. Figure 4 shows an example of the finished edge-bond.

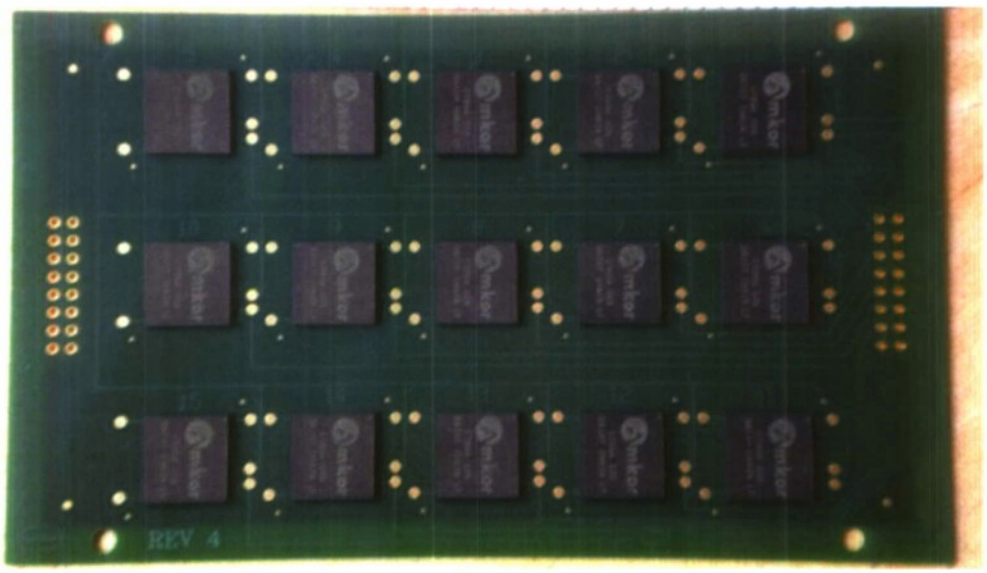


Figure 1. Test vehicle with components

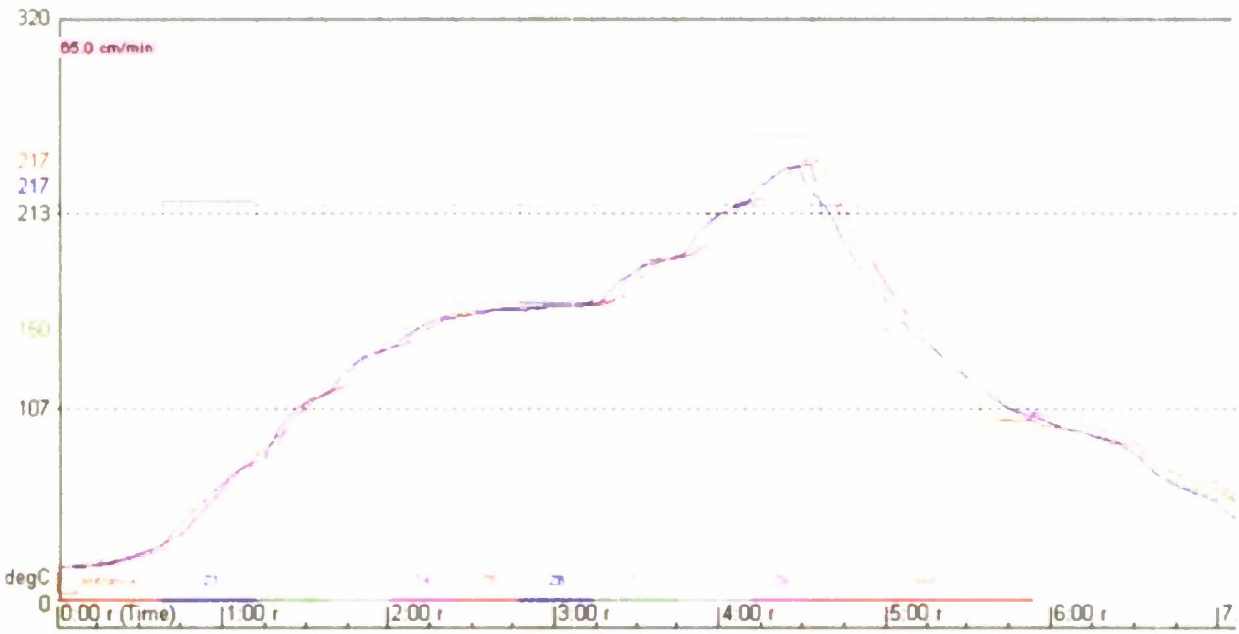


Figure 2. Solder reflow profile

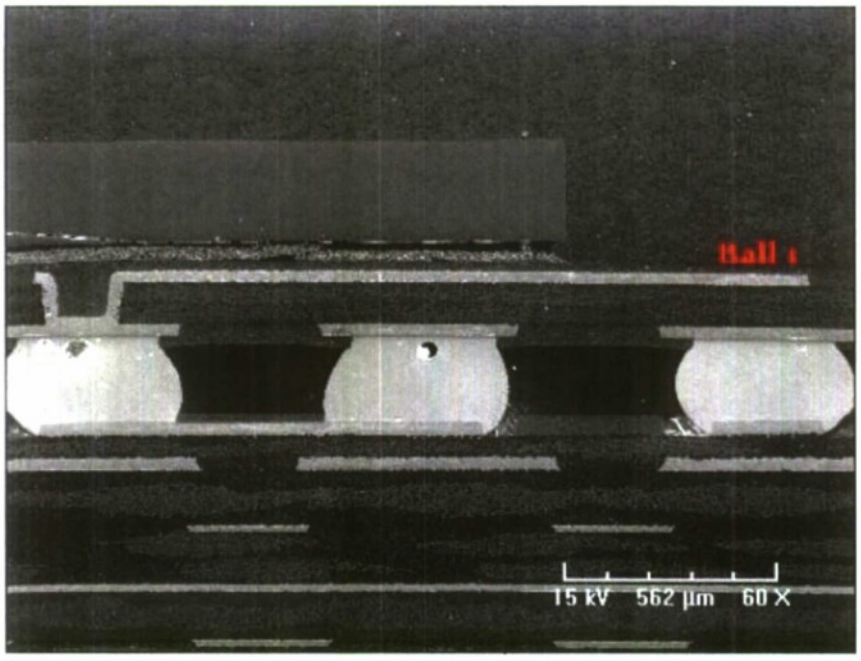


Figure 3. SEM of solder joints after assembly

Figure 4. An edge-bonded CSP

III. Drop Test Methodology

The drop tests were conducted using a Lansmont M23 TTSII shock test system, which applies a single half-sine shock impact pulse to the test vehicle for each drop. Multiple impacts were prevented by a pneumatic rebound brake system that slowly brings the table to a stop with minimal effect on board vibration. This shock test system represents a typical board level drop test setup.

For this study the test vehicle was mounted with the board in a horizontal position with the components facing downward which is the most severe orientation for board deflection [1, 19]. Four corner shoulder screws with 12 mm standoff supported the board mounted on the drop table. The drop table was secured between two guiding rods and could travel only along the vertical direction. When dropped from the chosen height, the drop table falls vertically and

impacts the stationary seismic shock mounted table base. This impact transfers an input acceleration pulse to the test board through the four corner supports. The acceleration peak value and pulse duration are controlled by the drop height, friction against guiding rods, and impact surface.

Three acceleration conditions were chosen from the JEDEC recommendations [2]: 900G, 1500G, and 2900G, with 0.7 ms, 0.5 ms, and 0.3 ms durations respectively. These are JEDEC condition F, B, and H. The input acceleration pulse was monitored for each drop by an accelerometer attached to the table base plate using a Test Partner TestPal signal conditioner and software. For each drop height and impact surface selected, the average result of two accelerometers was used as shown in Figure 5. The table impact surface varied between acceleration conditions, with a felt pad used for 900G and several sheets of watercolor paper used for 1500G and 2900G. The drop heights used were 368 mm (14.5 inch), 572 mm (22.5 inch), and 762 mm (30 inch). The drop height was adjusted incrementally to maintain consistent acceleration conditions during test cycles and the acceleration was measured on every drop. One deviation was made from the JEDEC standard in that the gap between the shoulder screw and board surface was controlled to within only 100 microns rather than the standard 50 microns [1]. A misalignment of tooling for the drop test support screws prevented use of the specified gap limitation.

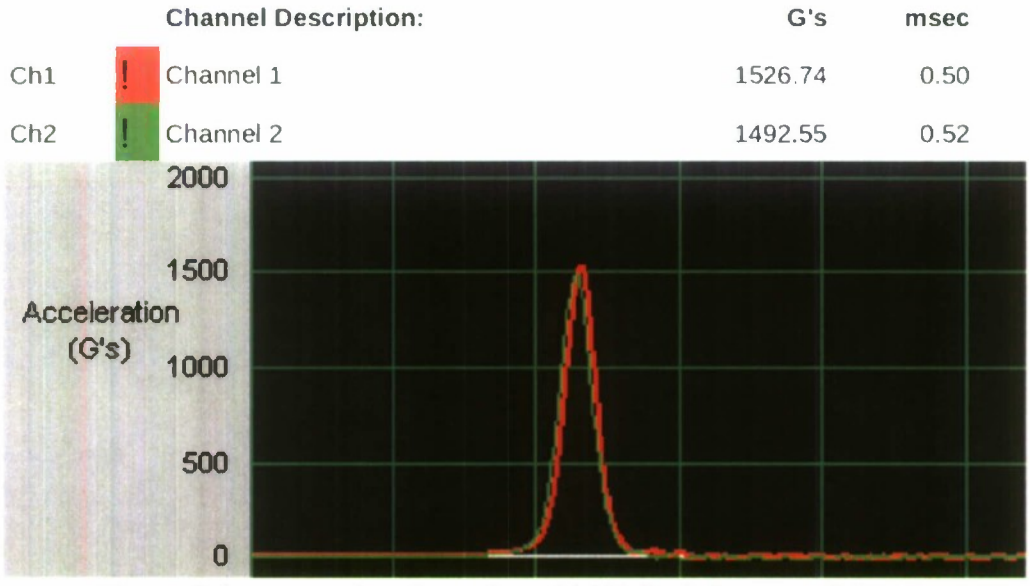


Figure 5. Input acceleration pulse of 1500G - 0.5ms, condition B

The test vehicles were split into two groups as shown in Table 1, one per failure detection system, with each group having 8 edge bonded boards (4 for each edge-bond material) and 6 boards without edge bond. Those groups were split so that each group would have 3 non-edge bonded boards at 900G, 3 non-edge bonded boards and 4 edge bonded boards at 1500G, and 4 edge bonded boards at 2900G. One additional board without edge bond was dropped at 2900G (no matching board was tested by post-drop).

Table 1. Number of boards per drop test variable eell

Failure Detection	DAQ System		Post-drop System	
Edge-bonding	Yes	No	Yes	No
900G	0	3	0	3
1500G	4	3	4	3
2900G	4	1	4	0

IV. Failure Detection Systems

This study compares drop impact failures with two failure detection systems: in-situ high-speed data aequisition (DAQ) with analog-to-digital conversion (ADC) yielding dynamic resistanee measurement, and post-drop statie resistanee measurement. These systems will be referred to as the data aequisition system and the post-drop system for purposes of diseussion.

The high-speed data aequisition system uses a National Instruments (NI) ADC, a desktop computer, and a voltage divider network to evaluate the resistance of the component daisy chain during the drop impaet at a sampling frequeney of 50KHz and 16-bit accuracy. This sampling rate of the system provides 50 sample points per milliseeond for eaeh component (50,000 per second), so that several samples are taken during the initial shock pulse (as short as 0.3 ms for 2900G). The primary deflection time of the board and first harmonie vibration frequency in a 1500G drop are near 4 ms and 240Hz [20]; with a 50Khz sampling frequeney this system provides more than 200 samples per board defleetion eyele. During each drop the ADC records every data point taken from the fifteen components and supply voltage and saves a data file for later analysis. A simple and proven method of aehieving dynamie daisy-chain resistance measurement at near real-time was used [21]. The daisy chain is plaeed in a DC series circuit with a statie resistor (R_s) of known value (in this ease 100Ω) to construct a voltage divider circuit as shown in Figure 6. The DAQ records the voltage (V_c), divided across the component resistanee and statie resistance. The voltage (V_c) relates to the resistance (R_c) by Eq. 1, where V_{DC} is the DC voltage souree set to 5V. As the component electrically fails, the resistance rises ($R_c \Rightarrow \infty$) and the DAQ registers a rise in voltage ($V_c \Rightarrow V_{DC} = 5V$).

$$R_C = \frac{V_c \cdot R_s}{V_{DC} - V_c} \qquad (1)$$

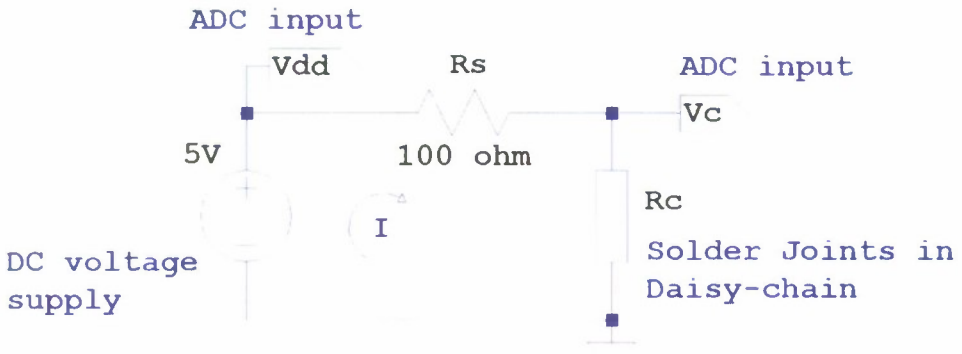


Figure 6. DC series voltage divider circuit

A cable was connected to the test vehicle by soldering the 16 wires (15 channels and common ground) directly into plated through-holes on the short side of the board. The cable was secured to the drop tester base plate to prevent loading against the solder connections during impact. The failure criterion used for the high-speed data acquisition system was taken directly from the JESD22-B111 standard [1]: a 100Ω resistance in the component daisy chain at any time during the drop impact or subsequent vibration is considered a failure, and the failure must be repeated in at least three out of five successive drops. Since a 100Ω static resistor and 5V DC supply voltage are used, the failure condition of 100Ω is the equivalent of measuring 2.5V on the component daisy chain. The electrical continuity of the cable-to-board through-hole solder joints was verified at regular intervals during and after drop testing to eliminate false positive failures due to broken cable connections.

The post-drop resistance measurement system used a LabView program and Keithley digital multimeter to read the daisy-chain resistance, once for each component after each drop, through a cable connected to the test vehicle after the vibration ceases. For this system no cable is connected to the board during the drop event. The failure criterion for the post-drop resistance measurement system of a static 10Ω rise (or more) from initial resistance is used.

The two failure detection systems use failure criteria that are necessarily different. The post-drop system is detecting a class of permanent failures only, which may be a solder crack that may be partially seated together when the board is at rest; the change in the resistance of the daisy chain due to this crack is small. The 10Ω static rise threshold was chosen to detect that small change. The data acquisition system detects intermittent failure, which may have insignificant resistance change when the board is at rest but a larger change during board deflection. The data acquisition system uses a temporary 100Ω resistance threshold, although data samples taken after the board vibration ceases could also be used to detect failure with the post-drop criteria.

The high-speed data acquisition system is capable of detecting intermittent failures as shown in Figure 7 during the board deflection and vibrations, whereas due to the single static measurement taken per drop the post-drop system can only detect permanent failures. In the example shown the static resistance rise from the initial condition is negligible and would not be detected as failure by the post-drop system since the static resistance rise is less than 0.1 ohm from the initial condition.

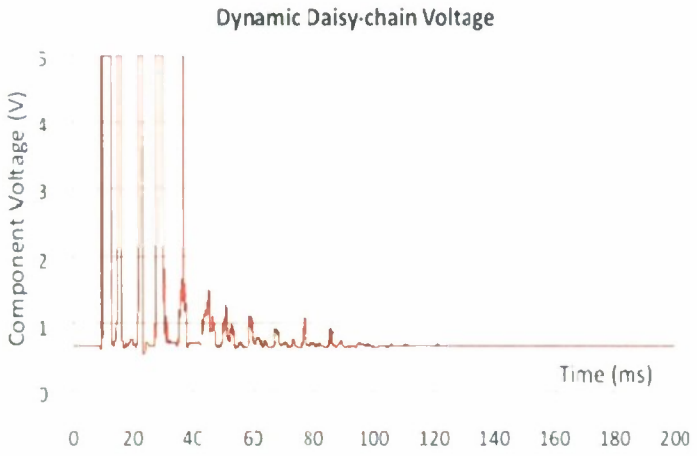


Figure 7. Intermittent failure detected by DAQ system, 10,000 data points shown in a 0.2 second window

V. Results and Discussion

Because the JEDEC standard drop test induces a complex strain pattern across the test board, causing varied stresses in the solder joints, JEDEC recommendations divide the components on the board into six groups (denoted A-F) that are expected to have similar failure rate due to the symmetry of their locations [1]. The issue of component location has been shown in a number of studies to be critical; the stress and strain in solder joints, and their failure rate, is partially dependent on the component location on the board [22, 23]. Che, et al. found that the maximum acceleration location occurs at the board center and is much higher than the input acceleration, however the maximum board strain occurs under components along the board edges and near the supports [22]. Therefore it is necessary to discuss failures in context of component location. The component locations are numbered as shown in Figure 8, and for this study the DAQ system cable is always soldered in thru-holes at the board edge near component 6.

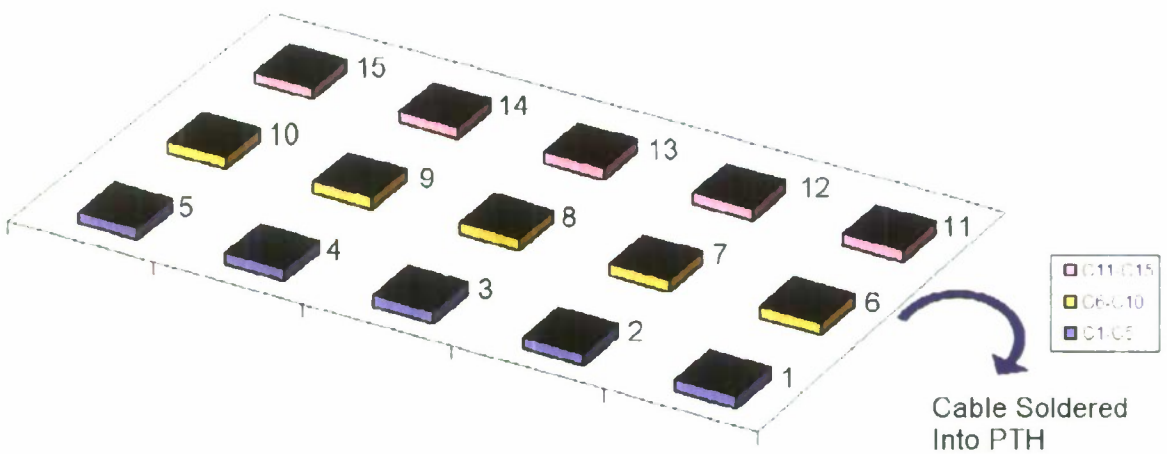


Figure 8. Board component location numbers for 15 components

A. CSP reliability without edge bonding

The drops to failure for each component location and test board without edge bonding are

shown in Table 2 for the DAQ system and in Table 3 for post-drop system. In both tables, each column represents one board except the first column. The first row is the input aeeeleration condition used for that board and the seeond row is the total number of drops the board was subjected to.

It is elear that the drops-to-failure is different between the DAQ system and the post-drop system. One of the most obvious differenees is that component C5 failed for every post-drop system board at 900G and 1500G but not at all for the DAQ system for those conditions. C14 and C9 also have similar differences between the deteetion systems, both failing more for the post-drop system than data aequisition. The reason for this differenee in failure between systems is not elear yet, however it may be due to the attaehment of the data acquisition eable to the board, whieh is believed to have an effect on board dynamies.

The eomponent location plays a signifieant role in the drop test reliability. The components along the board center (C3, C8, C13) tend to fail earliest and most frequently for both failure deteetion systems, although C3 did not fail as often for the DAQ system. Components C4 and C12 also failed eonsistently for both systems, but the symmetrieal board locations of C2 and C14 did not fail as often or as early; C2 did not fail consistently for either system, but C14 did fail for the post-drop system. The failure loeations have other symmetry issues as well; with C9 failing on the post-drop system but the symmetrieal mateh C7 failing infrequently.

It is interesting to note that the drops-to-failure vary signifieantly between different boards for the same component location. It is clear that higher G-level results in lower drops-to-failure. Every component exeept C10 in a board without edge-bonding failed after 50 drops when subjeeted to 2900G. Most of the eomponents fell off the board after less than 20 drops.

Table 2. Drops to failure by data aequisition, no edge bond

Aeeel (g)	900	900	900	1500	1500	1500	2900
Drops	75	75	100	70	40	60	50
Edge Bond	None	None	None	None	None	None	None
Component							
C1				**37**	**29**		**7**
C2							**25**
C3	**62**				**14**	**33**	**4**
C4	**26**	**26**	**34**	**26**	**6**	**23**	**4**
C5							**5**
C6					**21**	**35**	**3**
C7					**19**		**42**
C8	**28**	**44**		**50**	**3**	**13**	**7**
C9					**30**		**21**

C10							
C11					5		11
C12	16	6	43	13	2	6	4
C13	15	11	40	9	1	5	2
C14					21	32	38
C15							50

Table 3. Drops to failure by post-drop, no edge bond

Accel (g)	900	900	900	1500	1500	1500
Drops	75	70	100	70	40	60
Edge Bond	None	None	None	None	None	None
Component						
C1			82	55		38
C2						22
C3	7	31	15	8	3	11
C4	10	43	17	7	5	36
C5	65	2	14	1	5	14
C6	54					45
C7			61			9
C8	13	13	16	7	5	2
C9	53	16	11	28	8	14
C10						
C11	29		55			12
C12	6	9	18	5	3	3
C13	5	28	16	5	3	3
C14	1		37	5	34	4
C15	44		75	26		

B. CSP reliability with edge bonding

The drops to failure data for edge-bonded boards are reported in Tables 4 and 5. The total number of drops for each board is listed in row 2, and the edge bonding material (either thermal-cured epoxy or UV-cured Acrylated Urethane) is listed in row 3.

It is clear that edge-bonding improves the drop test reliability significantly by comparing the orange highlighted columns in Table 4 (2900G) to the last column of Table 2 (also 2900G). Eight components failed on a board without edge-bonding after 7 drops when subjected to 2900G as shown in Table 1, while first eight failures occurred for boards with edge-bonding after 36, 44, 100, and 133 drops when subjected to 2900G as shown in Table 4. For an input acceleration of 2900G, the edge-bonded boards show a 5-8 times reliability improvement.

The component location plays a significant role in the drop test reliability. Similarly to the boards without edge bonding, components C4 and C12 fail earlier than components C2 and C14, in the symmetrically mirrored board locations. Again as with the boards without edge bonding, components C7 and C9 show significant symmetry mismatch in both board failure

detection systems. This issue is explored further in the failure analysis section where the determined cause is explained.

The drop counts to failure are higher with edge bond applied for the majority of boards and component locations, and for both failure detection systems. However, the data acquisition system observed some intermittent failures that occurred for up to 150 consecutive drops in edge-bonded components without ever advancing to a permanent failure stage. In some of these cases the post-drop system would not have recorded failure when drop testing was stopped.

The data acquisition system does not always show failures in fewer drops (earlier detection) as was expected since it can detect the intermittent failure, but it recorded more total failures of the 2900G set than the post-drop system did. The capability of detecting failure earlier may be partially offset by the requirement of adding wired connections to the board during the drop impact; the wire may influence board deflection and vibration characteristics, and subtly effect drop reliability results.

Table 4. Drops to failure by data acquisition, edge-bonded

Accel (g)	1500	1500	1500	1500	2900	2900	2900	2900
Drops	325	350	279	355	190	170	175	173
Edge Bond	Heat	Heat	UV	UV	Heat	Heat	UV	UV
Component								
C1						151	66	61
C2		342	276		133	127		119
C3	80	292	33	101	70	72	12	103
C4	236	255	257		63	16		100
C5						36	73	91
C6		55				44	37	60
C7						35	69	158
C8	201			85	113	20	84	83
C9				292		25	29	124
C10			277			12	59	
C11		193	178	103		65	38	
C12	66	76	52	162	53	24	23	16
C13	61	129	73	77	42	13	18	14
C14		232				42	44	120
C15	107		268		44	22	25	90

Table 5. Drops to failure by post-drop, edge-bonded

Accel (g)	1500	1500	1500	1500	2900	2900	2900	2900
Drops	237	350	279	300	170	170	175	173
Edge Bond	Heat	Heat	UV	UV	Heat	Heat	UV	UV
Component								
C1		304	62			12	23	
C2			101				34	98
C3	2		180	81	74	72		23
C4	2	292	99	242		25	13	
C5	60		62	262		40		151
C6	112	282	180			151		
C7		6						
C8	88			108		68	30	21
C9		132		283	116	106	53	
C10		112						
C11	3	292				112		
C12	1	36	188	162	137	57	154	128
C13	159	99	188	133	6	144	36	43
C14	60			243			151	
C15				297				

VI. Failure Analysis

Failure analysis was performed on a subset of the failed test boards after drop tests. The outer row of solder joints of two components on two boards each was cross-sectioned. Scanning electron microcopy (SEM) images indicate the intermetallic layer thickness was 1-1.3 micron on the board side and 1.3-2 microns on the component side. To investigate the extent of cohesive failure resulting from the drop tests, the dye penetrant test was performed on eight boards, four with and four without edge-bond. Optical microscopy was used to identify dyed areas and determine failure location, root cause, and how widespread the under-pad resin cracking problem was for each component location.

A. Failure Modes

The most common failure observed was trace/pad breakage at the neck from the trace to pad as shown in Figure 9. The dyed area in the right of Figure 9 shows the resin cracked under the copper pad on the board side. The trace break was mainly due to the cohesive failure of resin between the copper pads and the fiberglass dielectric layer. Figures 10-12 show cross-sectioned solder joints where resin cracking is visible underneath the pads in the dielectric layer. The pad

cracking was commonly seen for both boards with and without edge bonding. Similar failure mode of trace broken has been reported by Chong, et al. [24], and the resin crack been observed by Mattila, et al. [16], Chong, et al. [24] and Wong, et al. [25].

All the components examined by dye penetrant that were electrically failed were categorized as solder failure, trace failure, both, or unknown. Of those components with electrical failure:

- 58% showed I/O trace failure with cracking under pads
- 12% showed solder joint fracture on the board side
- 19% showed both solder joint fracture and I/O trace failure
- 11% showed daisy chain trace failure and pad cracking

The 58% of electrical failures were due to one or two input/output trace connections broken away from the copper pad. Another 11% had pad cracking that led to daisy chain trace failure within the array (not at the I/O traces). The large ratio of electrical failures resulting from I/O traces cracked away from pads compared to solder joint fracture may be partially related to the test vehicle design and trace routing. Figure 13 shows a single CSP pad location with all four corners where corner 2 has two traces running outward from the component. These two traces are the daisy-chain input and output connections. Traces connected to the other three corners lead to test pads and are not part of the daisy chain. The orientation of every package on the test vehicle is the same, with corners 1 and 2 parallel to the short board axis, corners 2 and 3 parallel to the long board axis, and the orientation of the trace layout is also the same for each. Due to resin crack under the copper pad, the transition of the trace to the I/O pad is the weakest point, which causes copper trace/pad cracking. A copper pad crater and part of one of the corner 2 I/O traces is shown in Figure 14, which is the board side match of the failure shown in Figure 9. The frequency of this failure indicates that if the corner solder joint pads were allowed to lift off the board while maintaining electrical continuity (the daisy chained trace between solder pads may lift up), then the drop impact reliability of the assembly might be overestimated. A test vehicle utilizing typical PWB layout for CSPs in electronic devices, such as traces to vias rather than pad to pad daisy chains, may be more appropriate for evaluating board level drop impact reliability.

The secondary failure mode was solder joint fracture. Figure 15 shows a fracture near the board side Cu_6Sn_5 intermetallic layer. Solder fracture failures were observed at the board side only and no solder failures were found at the component side. Both complete and partial solder fractures were found by dye penetrant analysis as shown in Figure 16. It is interesting to note that both a solder joint fracture and a broken trace can lead to electrical failure as shown in Figure 17. The pad on the left side of the image in Figure 16 has resin cracking which led to trace breakage as the pad lifted away from the board with the component during board deflection.

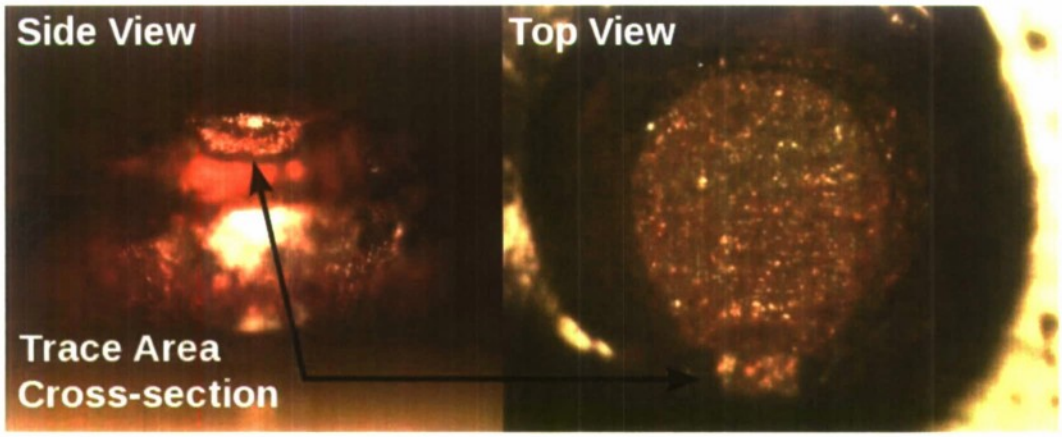


Figure 9. Trace cracked away from solder joint (left) and the same solder joint with pad dyed (view of component surface)

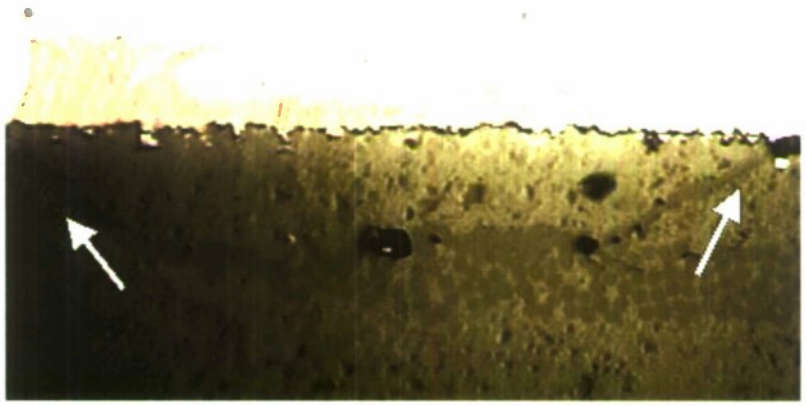


Figure 10. Cracked resin under the board side pad (dark line), edge bonded

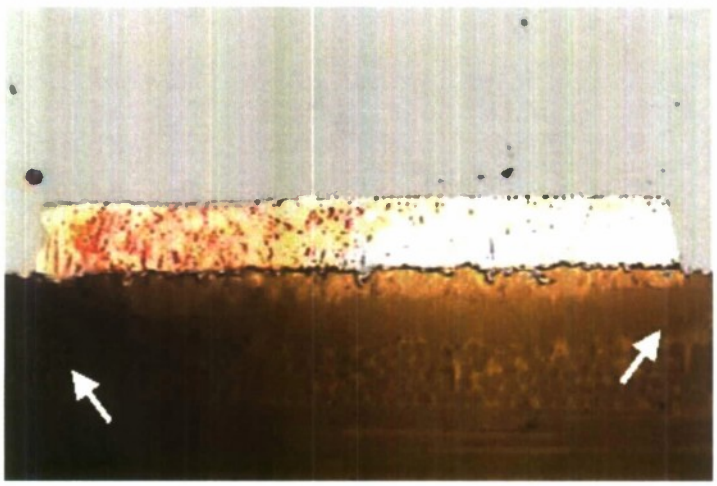


Figure 11. Crack in board resin underneath pad (thick dark area), no edge bond

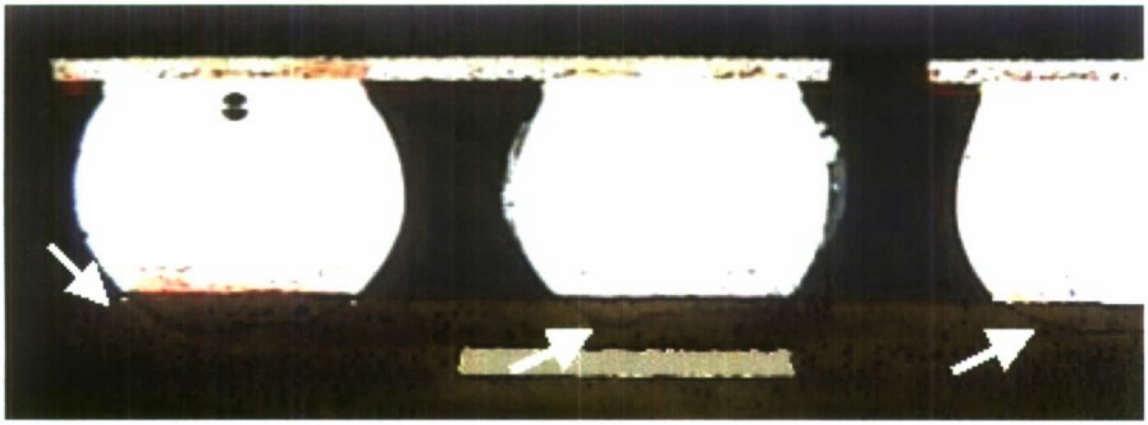


Figure 12. Cracked resin layer under pads for several solder joints, edge bonded

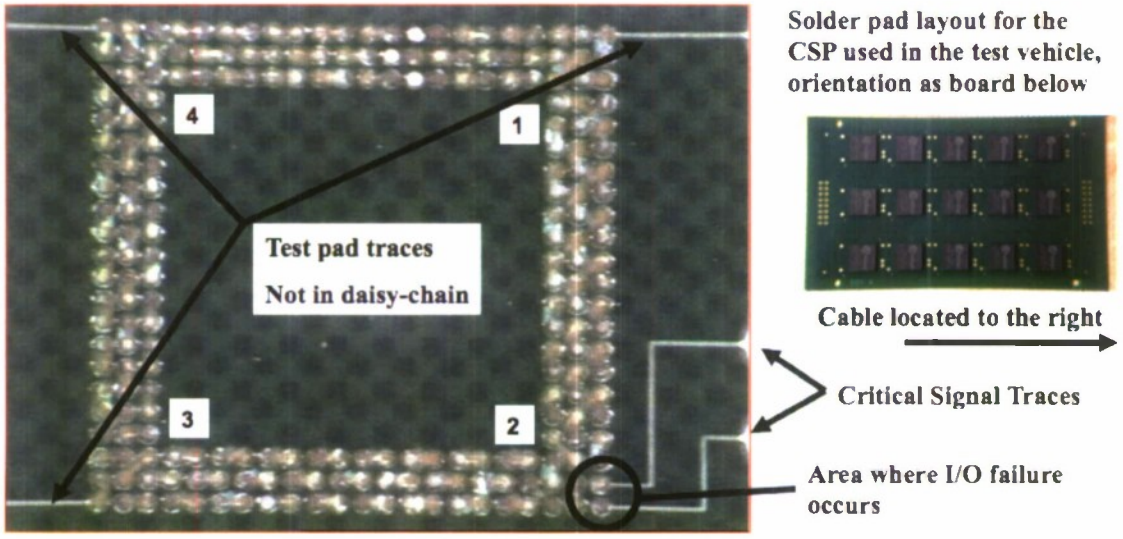


Figure 13. CSP I/O traces and component orientation

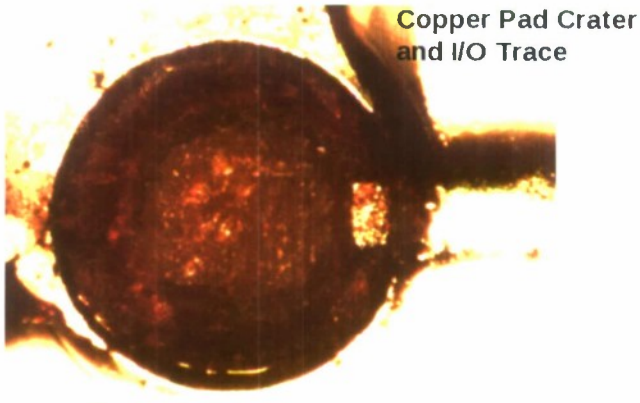


Figure 14. Copper pad crater with dyed board fibers

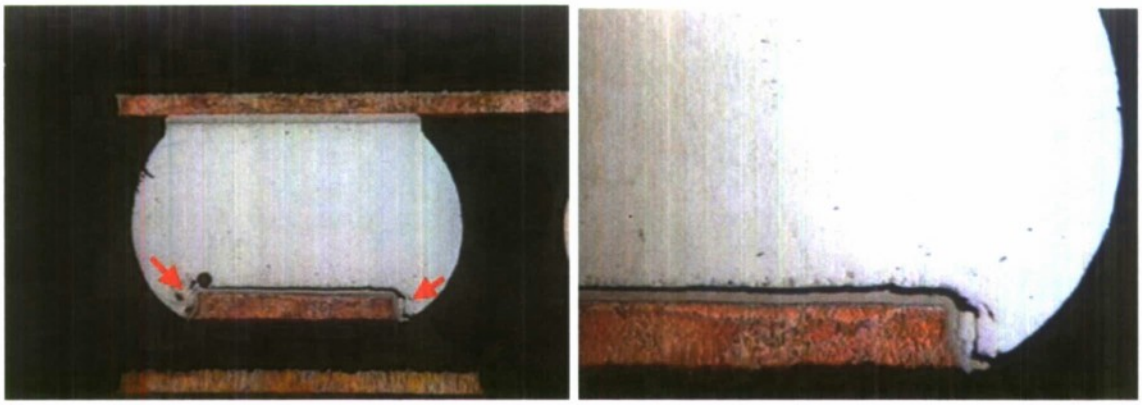


Figure 15. Solder joint fracture near the board-side IMC layer

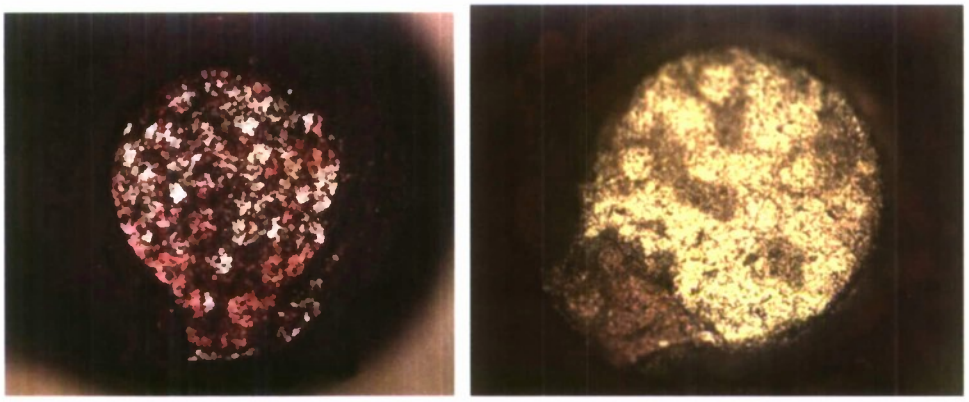


Figure 16. Complete fracture (left) and partial fracture (right)

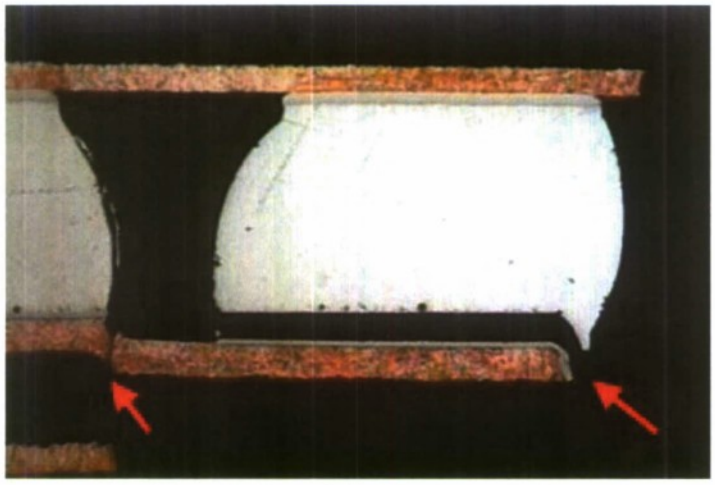


Figure 17. Board side solder fracture and broken trace

The extent of resin cracking under copper pad was examined for all components of the eight boards that were dye penetrant tested. Out of these 120 components (15 components per board times 8 boards), the relationship between electrical failure and resin cracking is summarized in Table 6. Table 6 shows that 72% of components were electrically failed and had resin cracking under the copper pad, while 19% of components were not electrically failed but

had resin cracking. The remaining 9% of components did not have resin cracking. This indicates that the solder joints are not the weakest link area of the assembly. It is recommended that board laminate materials be improved.

Table 6. Relationship between electrical failure & resin cracking

		Electrical failure	
		Yes	No
Resin cracking under pads	Yes	72%	19%
	No	6%	3%

B. Differences between edge bond material failures

There are notable differenees in the meehaniea1 failure mode between the two edge-bond materials. The epoxy material tends to fracture through the edge-bond material as shown in Figure 18. More than 20 components that were edge bonded with the epoxy material, or more than 10% of all the components in the group, dropped off the board during testing. This fracturing was observed to oeeur before eleetrieal failure happened. The aerylie edge-bond material did not fracture, but delaminated from the paekage sides. The aerylie was not observed to be delaminated from the board surfaee. Figure 19 shows that four undamaged edge bonds remained on the board after the component fell off. The properties of these two edge-bond materials are believed to contribute to the differenee in the meehaniea1 failure mode.

Figure 18. Fractured thermally eured epoxy edge bond

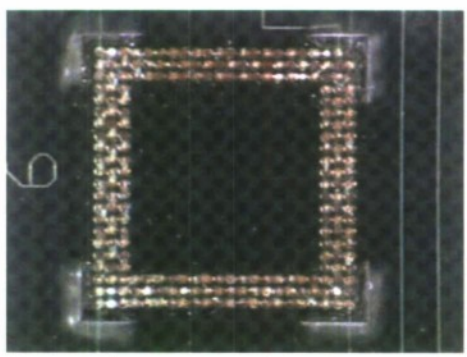


Figure 19. Four UV eured aerylie edge bonds on a board

VIII. Conclusions

The drop test reliability of 0.5mm pitch CSPs assembled on the JEDEC test vehicle with lead-free SAC305 solder are studied by two failure detection systems. The following conclusions can be drawn from this research:

1. The component location plays a significant role in the drop test reliability. Generally speaking, components at the center of the board are more prone to fail due to higher strains. But the differences in drops-to-failure between the two different failure detection systems indicate that additional mass on the board changed the distribution of strains along the board, which resulted in different failure locations.
2. Higher impact force or G-level resulted in lower drops-to-failure. But there are large variations in drops-to-failure between different boards under the same drop conditions.
3. Edge bonding can significantly improve drop test reliability. The edge bonded CSPs typically survived 5 – 8 times longer at 2900G-0.3ms drop impacts, and 8 – 10 times longer at 1500G-0.5ms drop impacts. But the performances of the two different edge-bond materials are different.
4. The majority of drop test failures were trace breaking caused by cohesive failure of resin between the copper pads and the fiberglass dielectric layer. This indicates that solder joints are not the weakest link area of the assembly. This pad cratering issue suggests that PCB laminate materials should be improved and the laminate specification and testing method should be included in a future JEDEC drop testing standard.

Acknowledgement

This work was supported by the Department of the Navy, Office of Naval Research, under Award #N00014-06-1-1111. The PI would like to thank students Andrew Farris for drop testing, Nicholas Vickers for failure analysis, Micah Deneeour for data organization, and Tom White for edge bonding. The PI would like to acknowledge strong support from Henkel Technologies and Flextronics International for both financial and technical support, especially from Dr. Dongkai Shangguan, Mr. Jasbir Bath, Mr. Dennis Willie, Mr. David A. Geiger, Dr. Brian J. Toleno, and Mr. Dan Maslyk.

References

1. JEDEC Standard JESD22-B111, "Board level drop test method of components for handheld electronic products", JEDEC Solid State Technology Assoc, 2003.
2. JEDEC Standard JESD22-B104-C, "Mechanical Shock", JEDEC Solid State Technology Assoc, 2004.
3. JEDEC Standard JESD22-B110A, "Subassembly Mechanical Shock", JEDEC Solid State Technology, 2004.
4. C.T. Lim and Y.J. Low, "Investigating the drop impact of portable electronic products", Proceedings of 52nd IEEE ECTC, 2002, pp. 1270-1274.
5. C.T. Lim, C.W. Ang, L.B. Tan, S.K.W. Seah, and E.H. Wong, "Drop impact survey of portable electronic products", Proceedings of IEEE Electronics Components and Technology Conference, 2003, pp. 113-120.
6. T. Y. Tee, H.S. Ng, C.T. Lim, E. Pek, and Z. Zhong, "Board level drop test and simulation of TFBGA packages for telecommunication applications", Proceedings of 53rd IEEE ECTC, 2003, pp. 121-129.

7. A. Syed, T.S. Kim, S.W. Cha, J. Scanlon, and C.G. Ryu, "Effect of Pb free alloy composition on drop/impact reliability of 0.4, 0.5 & 0.8mm pitch chip scale packages with NiAu pad finish", Proc. of 2007 IEEE ECTC, pp. 951-956.
8. A. Syed, T.S. Kim, Y.M. Cho, C.W. Kim, and M. Yoo, "Alloying effect of Ni, Co, and Sb in SAC solder for improved drop performance of chip scale packages with Cu OSP pad finish", Proc. of 2006 IEEE EPTC, pp. 404-411.
9. R.S. Pandher, B.G. Lewis, R. Vangaveti, and B. Singh, "Drop shock reliability of lead-free alloys -- Effect of micro-additives", Proceedings of 2007 IEEE Electronics Components and Technology Conference, pp. 669-676.
10. H. Kim, M. Zhang, C.M. Kumar, D. Suh, P. Liu, D. Kim, M. Xie, and Z. Wang, "Improved drop reliability performance with lead free solders of low Ag content and their failure modes", Proceedings of 2007 IEEE Electronics Components and Technology Conference, pp. 962-967.
11. S. Zhang, C. Chen, and S. Yee, "An investigation on the reliability of CSP solder joints with numerous underfill materials", Journal of SMT, Vol. 16, No. 3, 2003, pp. 25-30.
12. B. Toleno, D. Maslyk, and T. White, "Using underfills to enhance drop test reliability of Pb-free solder joints in advanced chip scale packages", Proceedings of 2007 SMTA Pan Pacific Symposium.
13. Y. Liu, G. Tian, R.W. Johnson, and L. Crane, "Lead-free chip scale packages: assembly and drop test reliability", IEEE Transactions on Electronics Packaging Manufacturing, Vol. 29, No. 1, Jan. 2006, pp. 1-9.
14. G. Tian, Y. Liu, R.W. Johnson, P. Lall, M. Palmer, M.N. Islam, and L. Crane, "Corner bonding of CSPs: processing and reliability", IEEE Trans. on Electronics Packaging Manufacturing, Vol. 28, No. 3, July 2005, pp. 231-240.
15. H. Lou, X. Qu, Z. Chen, J. Wang, T. Lee, and H. Wang, "Lifetime assessment of solder joints of BGA package in board level drop test", Proc. of IEEE 6th International Conf. on Electronic Packaging Technology, 2005, pp. 696-700.
16. T.T. Mattila, P. Marjamaki, J.K. Kivilahti, "Reliability of CSP interconnections under mechanical shock loading conditions", IEEE Trans. on Components and Packaging Technologies, Vol. 29, No. 4, Dec. 2006, pp. 787-795.
17. D.Y.R. Chong, K. Ng, J.Y.N. Tan, P.T.H. Low, J.H.L. Pang, F.X. Che, B.S. Xiong, L.H. Xu, "Drop impact reliability testing for lead-free & leaded soldered IC packages", Proceedings of 2005 IEEE Electronics Components and Technology Conference, pp. 622-629.
18. T.Y. Tee, H.S. Ng, C.T. Lim, E. Pek, "Drop test and impact life prediction model for QFN packages", Journal of SMT, Vol. 16, No. 3, 2003, pp. 31-39.
19. Y.S. Lai, P.F. Yang, and C.L. Yeh, "Experimental studies of board-level reliability of chip-scale packages subjected to JEDEC drop test condition", Microelectronics Reliability, Vol. 46, No. 2, 2006, pp. 645-650.
20. P. Marjamaki, T. Mattila, J. Kivilahti, "Finite element analysis of lead-free drop test boards", Proceedings of 2005 IEEE Electronics Components and Technology Conference, pp. 462-466.
21. J.E Luan, T.Y. Tee, E. Pek, C.T. Lim, Z. Zhong, J. Zhou, "Advanced numerical and experimental techniques for analysis of dynamic responses and solder joint reliability during drop impact", IEEE Transactions on Component and Packaging Technologies, Vol. 29, No. 3, Sept. 2006, pp. 449 – 456.

22. F.X. Che, J.H.L. Pang, W.H. Zhu, W. Sun, A. Sun, C.K. Wang, H.B. Tan, "Comprehensive modeling of stress-strain behavior for lead-free solder joints under board-level drop impact loading eonditions", Proeeedings of 2007 IEEE ECTC, pp. 528-534.
23. S. Park, C. Shah, J. Kwak, C. Jang, J. Pitarresi, T. Park, S. Jang, "Transient dynamic simulation and full-field test validation for a slim-PCB of mobile phone under drop/impact", Proceedings of 2007 IEEE ECTC, pp.914-923.
24. D.Y.R. Chong, F.X Che, J.H.L. Pang, L. Xu, B.S. Xiong, H.J. Toh, and B.K. Lim, "Evaluation of influeneing factors of board-level drop reliability for chip scale packages (fine-piteh ball grid array)", IEEE Transaetions on Advaneed Paekaging, 2007.
25. E.H. Wong, R. Rajoo, Y.M. Mai, S.K.W. Seah, K.T. Tsai, and L.M. Yap, "Drop Impact: Funamentals & Impaet Charaeterization of Solder Joints," Proceedings of 2005 IEEE ECTC, pp. 1202 – 1208.

HONORS PROJECT

BalloonSAT:
Low Altitude, Low Cost Remote Sensing System

Project Investigators:

Dan Cary, Electrical Engineering
Thomas Fulenwider, Electrical Engineering
Carlos Hurtado, Aeronautical Engineering
(John Saghri-Advisor)
California Polytechnic State University
San Luis Obispo, CA

Aerial Platform and Stability

Introduction

The BalloonSat system is comprised of a tethered helium-filled weather balloon and the aerial imagery platform. The helium balloon was filled to about 8 feet in diameter and secured to the ground by four tethered lines. At the actual flight test, people were employed to hold down the tethered lines. The aerial imagery platform, or "payload", consists of an upper box, which stores the electronics and batteries, a C-bracket which provided the stability for the panning motion, and a lower box which housed the camera, infrared lens and changing mechanism, and also a shutter release servo and mechanism. A fishing net was fixed around the balloon to support the payload and prevent the entire project from floating away.

Design and Construction

The design was divided into three different parts: the upper box, C-bracket, and the lower box. The upper box was constructed out of aluminum due to its low cost, excellent workability, and light weight. The box was made out of one piece of aluminum sheet metal and the sides were bent up to create a box. The side supports were bent using the same bending machine in the workshop at the Cal Poly campus and were spot welded onto into place on the box. The lid was also made out of aluminum. The upper box housed the batteries, transmitter, circuit board, panning servo, shaft, bearing system, and the circuit board. Figure 1 depicts the schematic of the upper box.

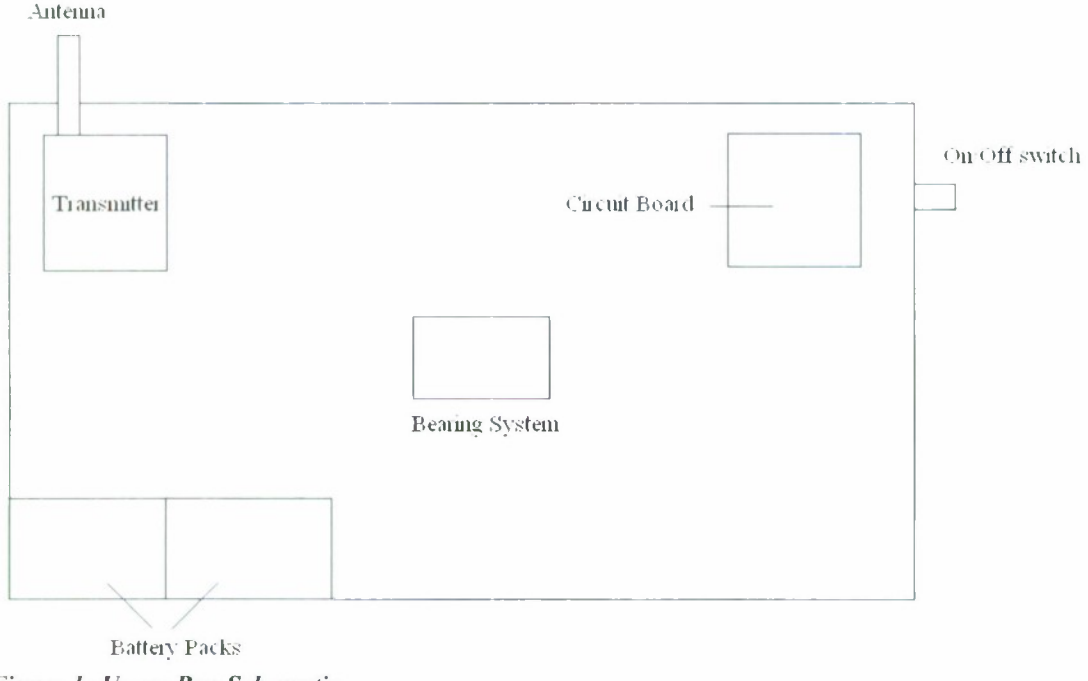


Figure 1: Upper Box Schematic

The upper box lid was going to be held in place with bolts, but it was decided to use 10" bungee cables instead. Figure 2 illustrates the outside of the upper box.

Figure 2: Upper Box

The shaft used to support the lower box is attached at one end to the servo at the top of the upper box, then extends through the box and down into the C-bracket.

The C-bracket is made out of sheet metal as well. Its dimensions are 6.25" L x 4" W x 6" H. It was designed and built to be able to support the weight of the lower box and serve as the support for the panning servo. Figure 3 depicts the C-bracket on BalloonSat.

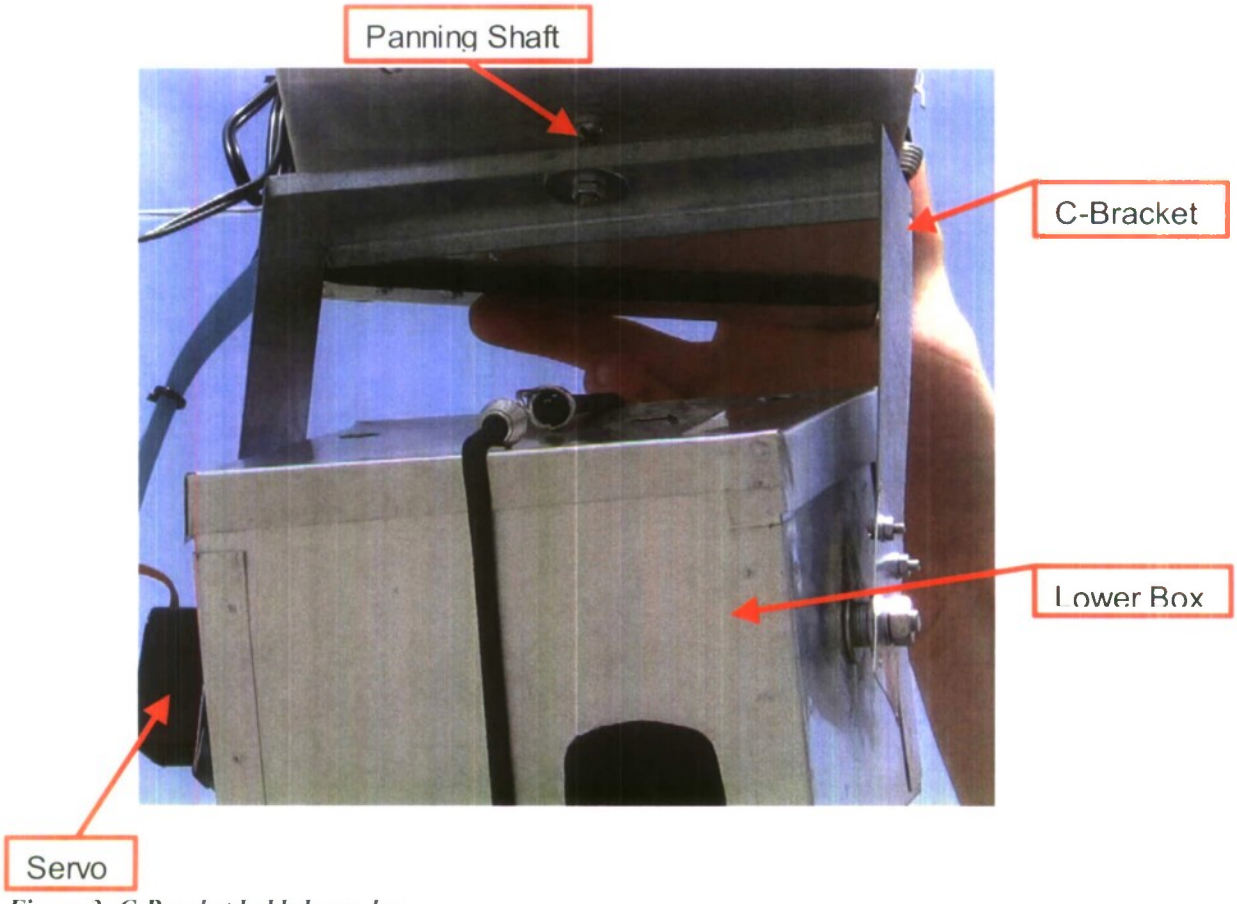


Figure 3: C-Bracket holds lower box

It was determined that a C-bend be used on the C-bracket. This added better compressive, torsional, and buckling performance onto the C-bracket. As a result of this bend, the C-bracket was much stiffer and more structurally sound. One end of the C-bracket housed the tilting servo. The other end provided the fixed point for the lower box. This allowed the lower box to rotate in a tilting manner, allowing another degree of freedom for our lower box and making it much more flexible in choosing the particular orientation of our pictures.

The lower box is constructed out of aluminum also and its main purpose is to house the camera, shutter release servo and mechanism, and the infrared filter and mechanism. The lower box measures 6" L x 5" W x 4.5" H. Figure 4 illustrates the configuration of the lower box.

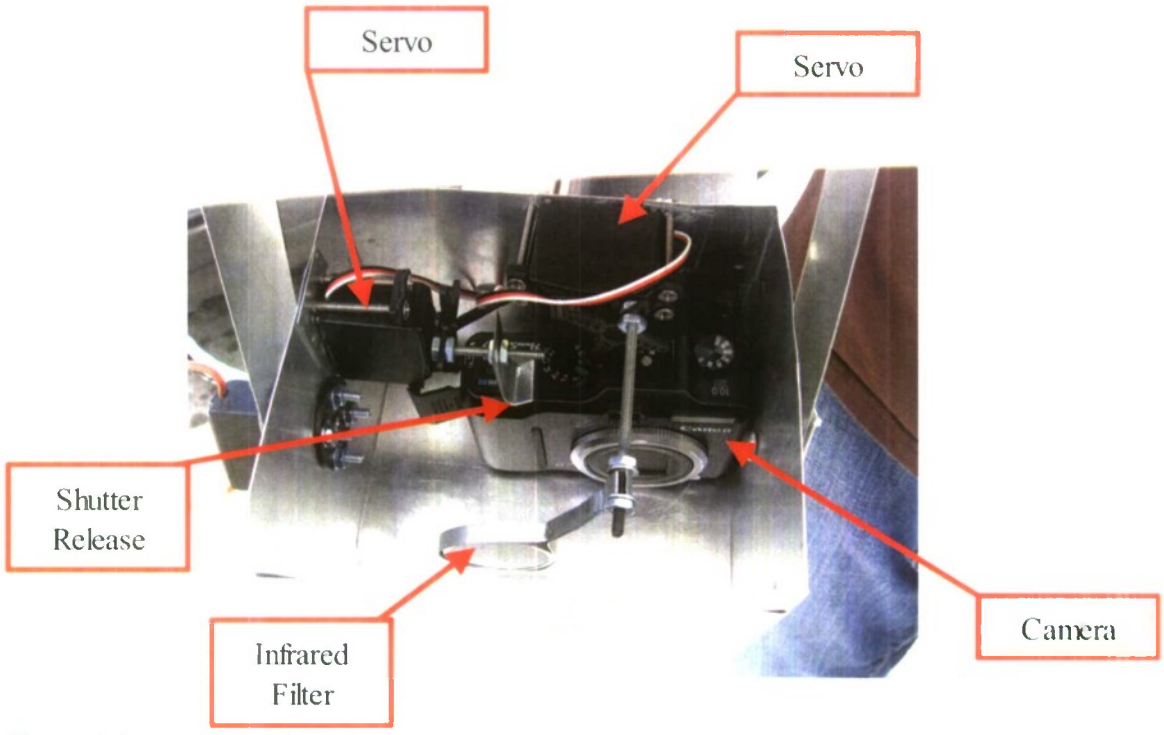


Figure 4: Lower box configuration

Similar to the upper box, the lower box was constructed from a single piece of aluminum. The sides were bent upward using a bending machine at the workshop, then reinforced with a piece of aluminum that was spot welded to hold two different sides together. Four corner supports were added to provide extra reinforcement. These supports were spot welded in three different spots on each support. The tilting servo was mounted onto one side and the other side utilized a support screw to allow the lower box to rotate freely.

The servos used for the infrared mechanism and the shutter release were mounted onto the lower box. Holes were carefully drilled so that 1.5 inch bolts could be fastened to hold up the servos. The camera was secured in place by it's built-in tripod mount, using a bolt through the box floor and washers.

The infrared filter changing mechanism was composed of a servo, shaft, and the infrared filter housing. The shaft was about 3/16 of an inch in diameter. The infrared lens housing was

constructed using a very ad hoc approach. A piece of sheet metal was cut to the same width as the infrared lens and a length that would be long enough to wrap around the lens, letting it hang from the servo shaft. The housing had to be hand-bent for the safety and security of the brittle lens. The loop that was made to fit around the servo shaft was spot welded at several points. This filter housing was secured with nuts onto the threaded shaft.

The shutter release mechanism was also made out of sheet metal. Basically, a square was cut out and bent to provide enough force to activate the camera's shutter. Then, a hole was drilled onto it so that the threaded shaft could be inserted and secured using washers and nuts. A larger hole had to be drilled onto the servo head, similarly to the infrared servo. Figure 5 shows what the completely integrated BalloonSat system looked like.

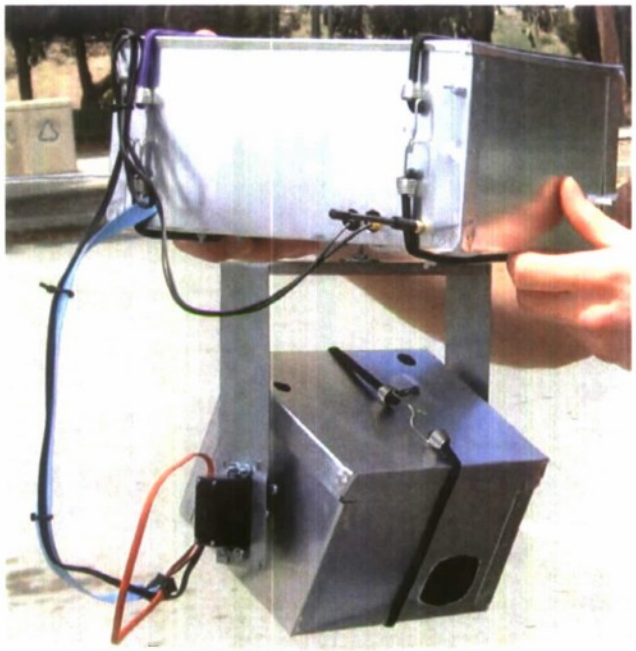


Figure 5: Integrated BalloonSat

We purchased gray fishing net from an online source, crisscrossing it over the balloon and securing the spaces together with fishing line to make sure the balloon did not drift away.

A helium balloon was ordered from Edmund's Scientific. We decided that inflating it to 12 of the possible 19 feet would provide a necessary lifting capacity of over 56 pounds. However, due to a lack of helium, we were only able to inflate our balloon to a diameter of about 8 feet, providing a lifting capacity of about 16.6 pounds. We knew this would not be enough because our payload weighs about 12 pounds, our balloon weighed about 3 pounds, and our fishing net weighed less than a pound. We had no other choice but to launch the partially filled balloon and hope for the best.

As expected, our weather balloon did not go up as high, and the high winds required that the tethered lines be secured on a flexible platform, meaning that we used our friends as the ground supports. This actually worked out better, especially in avoiding trees. Our balloon reached a

maximum altitude of about 200 feet, which was well short of our target height of 1000 feet. A total of about 160 pictures were taken, or 80 pairs of colored and infrared images.

BUDGET

Our program budget was $1500. We are proud to announce that we stayed within that budget. Table 1 shows how the money was used.

Table 1: BalloonSat budget.

Item	Cost ($)
Electronics	
Camera	639.86
Video Transmitter	57.99
Hobby RC Set-up	21.22
Components	30.34
Servos	30.99
Batteries	34.88
Structure	
Construction Materials	34.16
Jute netting	53.86
Helium	107.75
19' and 8' weather balloons	132.98
Mini bungees	3.49
Stakes	3.23
Item	Cost ($)
Fishing net	50.00
Open House Display	
Kinko's prints	28.36
Poster print	20.00
Laminating job	3.99
Wall signs and construction paper	21.56
Display board and tape	16.13
TOTAL:	**$1,290.79**

The most expensive item in our project was the camera. We performed a trade study at the beginning of the project, and felt the Canon G7 was the best choice, based on cost, weight, shutter speed, mega pixels, and ISO.

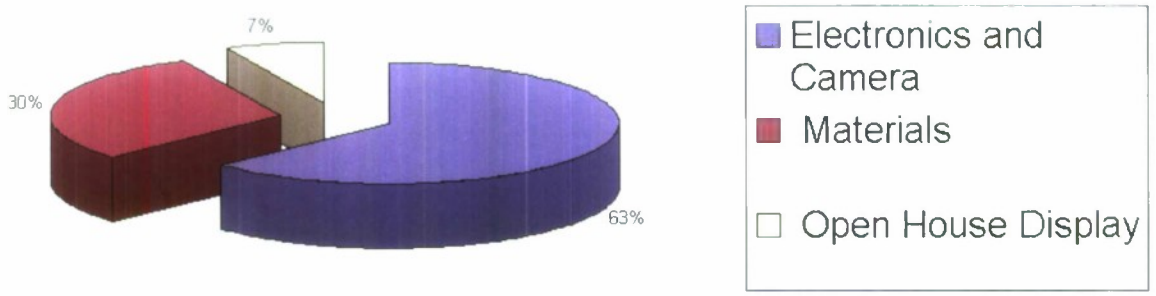


Figure 6: Dynamic view of BalloonSat Budget

Lessons Learned and Advice

As with any engineering project, several aspects of the BalloonSat system can be improved on. The following is a list of some suggested improvements for future BalloonSat missions.

- Buy the weather balloon far in advance. This would also include a close inspection of older balloons that might be flight material.
- Bring 2 tanks of helium. This is crucial because one tank was only enough to inflate the balloon to about 8 feet in diameter.
- Do not buy items that you will not use. This was especially true of our jute netting that we later discarded because it was too heavy.
- Complete more of the work the first quarter. This could even include beginning the manufacturing the housing. This way, you are not rushing to get this done the last few weeks of spring.
- There is no need to completely re-invent the wheel. The camera which we purchased, the Canon G7, is a very capable camera and will not need to be replaced within a few years. It is 10 mega pixels, and very lightweight. Also, future teams can use it again to keep cost low. Almost half of our budget was spent on this camera. It also has a 2 GB memory card, so it does have the capacity for several pictures. Also, our structural housings can be reused on future missions.

Electronics

Introduction

The real theme behind the electronics for the our balloon was "off-the-shelf and inexpensive." We had limited time to get many systems working together, so utilizing off-the-shelf components ensured plug-and-play simplicity when it came to assembling the components. Since much of our budget went toward purchasing our camera, the supporting electronics needed to be cheap so we could maintain our "low-cost" goal.

Methods

In order to organize the goals and capabilities of the electronics for the BalloonSat, we created a generalized block diagram for both the onboard and ground station electronics. These block diagrams can be seen in Appendix A. The diagrams offer a very high level, or broad perspective of the electronics system as a whole. From them we were able to construct each component of the system separately while keeping tabs on how each component needed to interact with others.

The next step of the project was to find parts that would enable the balloon to perform all the tasks that we required of it. Appendix B shows the electronics parts list. The R/C transmitter and receiver enabled us to remotely control the operation of the balloon from the ground station. The video transmitter and receiver offered visual feedback, letting us see what the camera was seeing.

Controller Board

The next key step was to start designing and fabricating a controller board. This board manages power and controls everything onboard the balloon. The schematic for the control circuit was made first using the freeware design software, Eagle Layout Editor. This software offers a simple user interface that combines both schematic capture and board layout.

The circuit was designed in order to accomplish all the capabilities that were outlined in the system block diagrams. One important addition to the circuit was a 12 volt regulator that could be turned on and off from the microcontroller. This regulator supplies power to the video transmitter, which happens to be somewhat of a power hog. The ability to turn this transmitter on and off gives us the option to go into a low power mode to conserve our limited power. See Figure 7 for the schematic.

After the schematic was built and tested physically on a breadboard, it was exported to a board layout. The schematic is a way to represent a circuit through simple symbolic picture whereas the board layout represents the circuit through the actual physical components and their physical interconnections. Some of the goals for the board was to keep it under two inches square, minimize the connections on the top side of the board, and to include component labeling on the top side of the board. These goals were met, as seen in Figure 8.

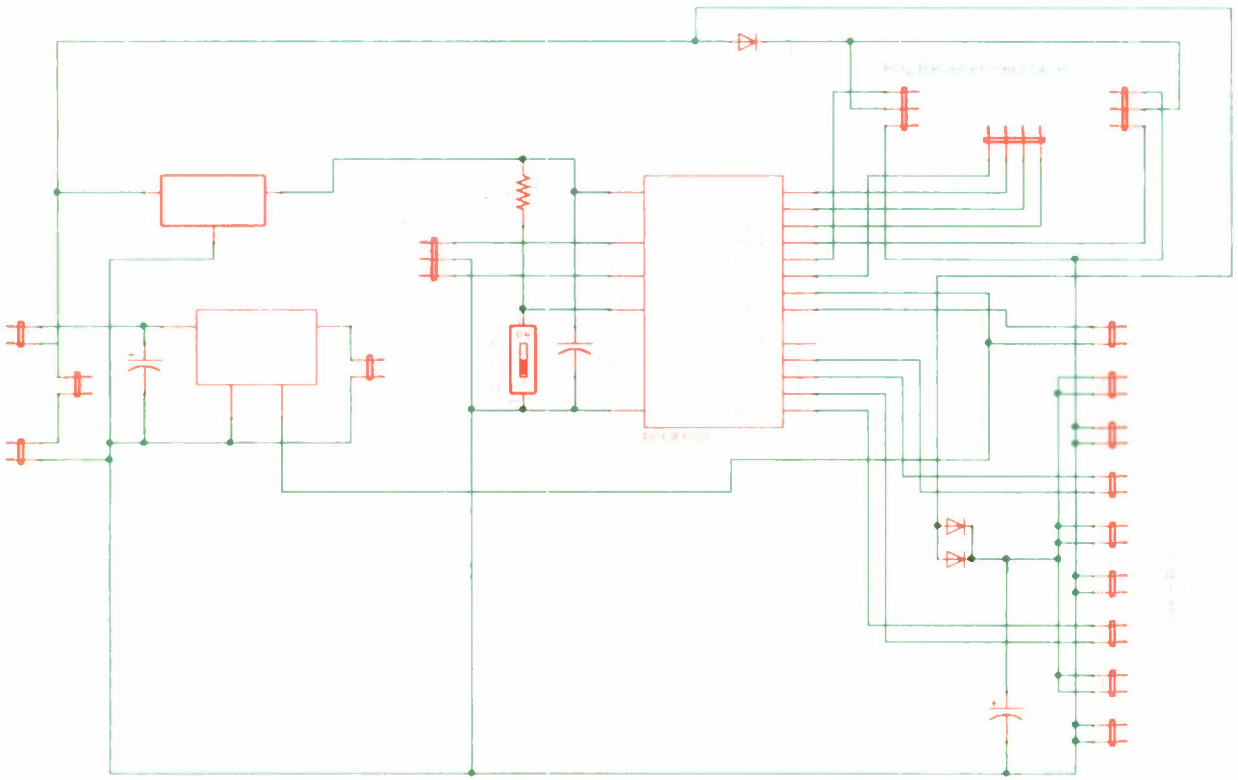


Figure 7: BalloonSat Controller Schematic

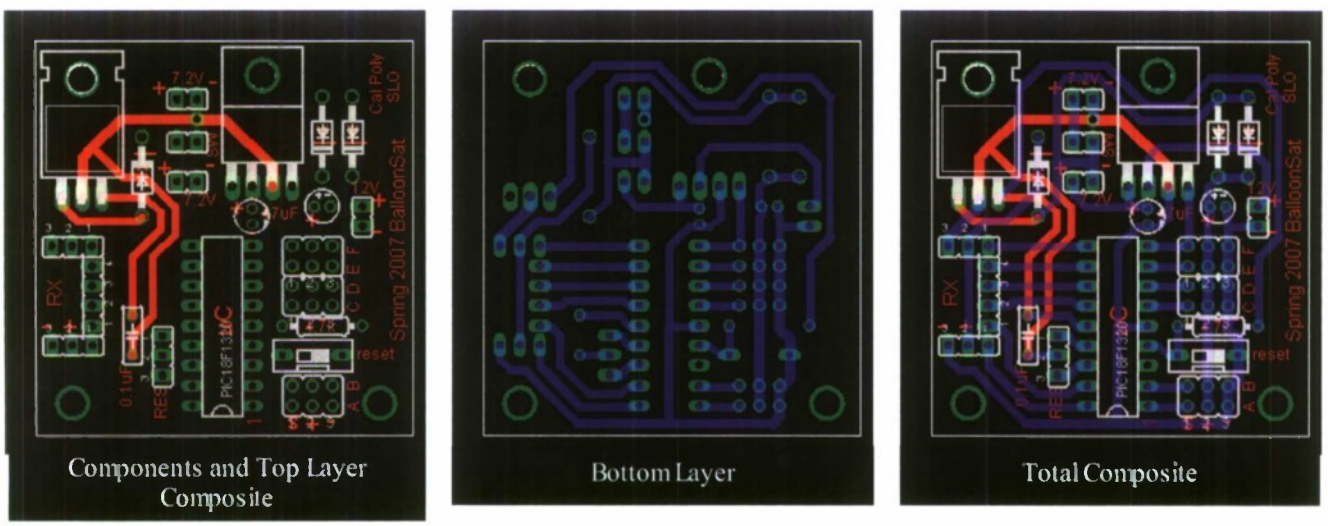


Figure 8: BalloonSat Board Layout

This board was fabricated from scratch, which was a challenge but saved us money. The process used to fabricate the board was the Toner Transfer Method. This method uses readily available tools and resources to create quality, professional-looking circuit boards. You start with a fiberglass board completely covered on both sides by copper. Unwanted copper is then removed, leaving behind copper traces that form the interconnections between circuit components. The process is as follows:

1. *Print:*

   Each layer of the board must be converted to a black and white image. Anywhere that copper is to remain must be black and anywhere that copper is to be removed must be white. As seen in Figure 9, the component labeling was made out of copper, so all the lettering appears black.

   One important thing to remember is that the top layer must be mirrored. If it is not, then the image will be inverted once it is transferred onto the copper clad fiberglass.

   The layers are printed out on a laser printer using photo paper. Photo paper is used so that the black toner from the printer does not sink into the paper but rather sits on the glossy surface. An inkjet printer will not work since the ink cannot be removed from the paper like toner can.

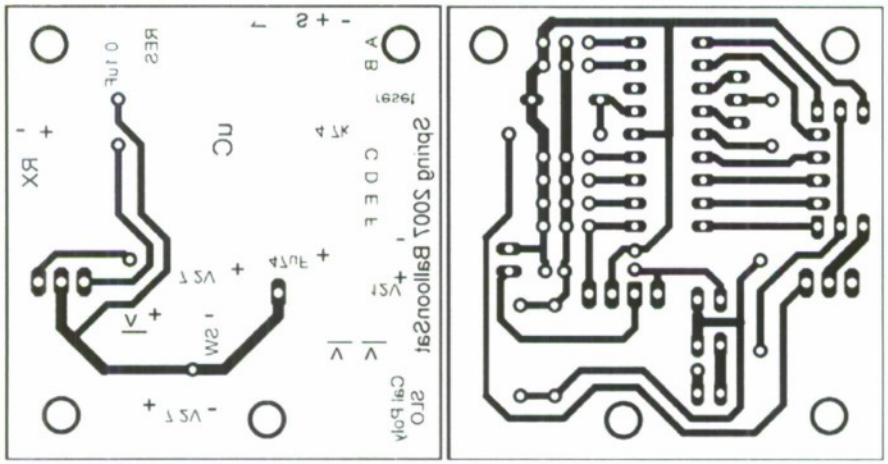


Figure 9: Top and Bottom Layers to be Printed

2. *Iron:*

   Both printed layers are then positioned on each side of the cleaned copper clad board. The toner must be facing the copper so that it can be melted to the copper. The layers must be perfectly aligned or else the holes that are drilled will not be consistent with the top and bottom layers. The paper is held in place with tape while the toner is melted to the copper using an iron on medium heat and no steam. It is crucial to move the iron around and apply pressure across the entire board so that every bit of toner is transferred.

3. *Peel:*

   Once both sides of the board have been ironed thoroughly, the paper must be removed. It is best to soak the board and paper in water for a couple of hours before attempting to remove the paper. The photo paper is carefully peeled away, leaving the toner on the copper. It is important that all the paper is removed because if it is not, copper that was suppose to be removed may not be, leaving a potential for electrical shorts in the circuit.

4. *Etch:*

The etchant used for this project was ferric chloride, although there are many other etchants that could have been used. Whichever one is used, it is important to note some safety precautions regarding their use. Etchants are designed to eat metal, so do not use any metal utensils or perform this step near any metal that you would like to preserve. Use only plastic containers and utensils. The chemical reaction to remove the copper from the board releases small amounts of harmful gas, in this case chlorine, so operate in a well ventilated area. Since the etchant is a very strong acid, wear latex gloves, pants, and close toed shoes. This is not a dangerous process, but safety should always be first.

Ferric chloride generally comes in a crystalline solid form, so it has to be mixed with water before it can be used. Ferric Chloride is slowly added to a plastic container filled with warm water to a half an inch deep. Once the correct concentration of acidity is reached, the copper clad board with its black toner traces is set in the acid.

The container must be gently agitated in a circular motion so that fresh acid is always flowing over the board. It is also a good idea to flip the board occasionally in order to ensure that the top and bottom layers are etching evenly. The etching process took about 20 minutes using ferric chloride. For the first 15 minutes nothing appears to be happening as the acid eats through the copper. During the last five minutes, holes begin to appear in the copper and the progress is easier to track. The etching time can be reduced by using a different etchant, using a heater to keep the acid warm, or by using a non-metallic pump to circulate the acid more efficiently.

5. *Clean:*

The board should be cleaned in water to remove any stray acid. The board is then cleaned with acetone, which removes toner and reveals the copper traces.

6. *Drill:*

Mini drill bits are used to drill the appropriate holes in the fiberglass. Be sure to wear protective eye wear since these mini drill bits are very susceptible to breakage. A small hammer and awl are also quite useful for marking the holes before drilling them. For a view of the controller board after etching and drilling, see Figure 10.

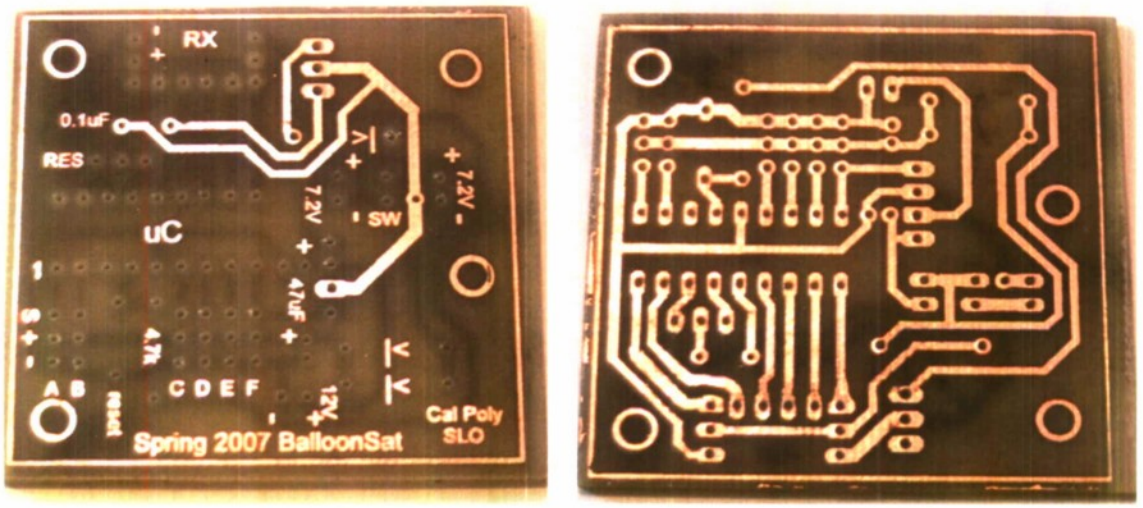
7. *Lacquer:*

The board now consists of bare copper traces laid out over a fiberglass board. The board needs to be lacquered in order to protect the integrity of the circuit. The lacquer insulates the copper so that adjacent copper traces cannot be electrically shorted. The lacquer also protects the copper from oxidizing, which would reduce its conductivity and degrade the circuit's performance.

8. *Solder:*

Solder is a type of metal that is used to make both electrical and mechanical connections from the controller. Each component, such as the microcontroller or voltage regulator, is

inserted into the appropriate holes in the controller board and then soldered into place. The solder bonds the component to the copper traces on the board in order to complete the circuit. The partially soldered board can be seen in Figure 11.



Top Layer Bottom Layer

Figure 10: Control Board after drilling

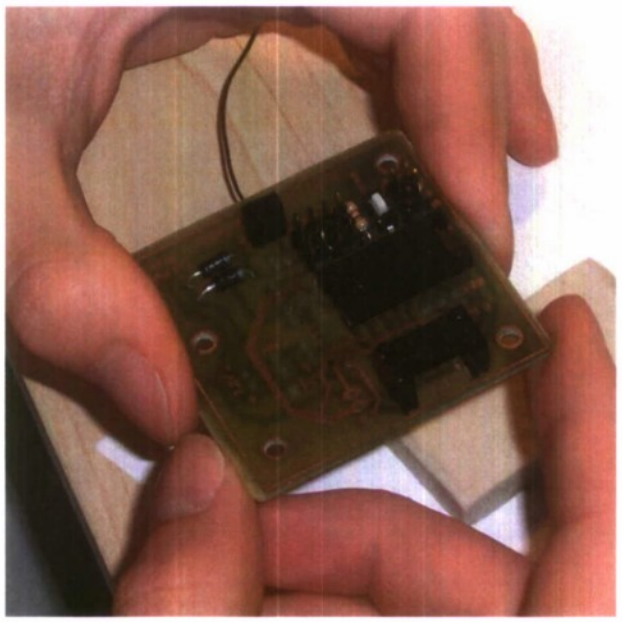


Figure 11: Control Board after soldering some components

Programming the Microcontroller:

The microcontroller had to be programmed to perform the tasks that we outlined in the system overview block diagrams. Since there was no complex processing that had to be done, a simple high-level compiler was used to create the code that would determine the microcontroller's operation. This code, written in Mikroelektronika's MikroBasic, can be found in Appendix C. After programming, the microcontroller can read the R/C receiver for commands, move the

camera using the pan and tilt servos, and take pictures. The picture taking process is set up to take a color image, move the infrared filter in front of the camera, and then take an infrared image, all with the push of a single button on the R/C transmitter.

Challenges

The electronics were closely related to the physical structure of our BalloonSat, so much of the structure was dictated by the needs of the control system. Since the operation of the electronics was dependent on the structure itself, the electronics could not be fine-tuned until the structure was complete. This put a major time crunch on getting the controls finished once the structure was finally completed.

The structure itself makes it very difficult to test the controls and to modify anything. The lower box must be detached before anything in the upper box (controller board, batteries, pan servo, and video transmitter) can be accessed. Because of this, the majority of testing with the controller board had to be done with the structure partially disassembled. Changing camera settings was also very difficult since the only way to get the camera out was open the lower box, unscrew the camera from the bottom, wiggle the camera until is was clear of the servos, unplug the video cable, and sneak the camera up between the servos and the infrared filter. Needless to say, a more accessible design is desirable.

The servos used to operate the shutter release and the filter rotation were originally suppose to be mini servos, but they were never shipped. Standard sized servos had to be used in their place. The standard servos were far too powerful for the task. Since the filter rotation shaft did not run through some sort of bearing in the front of the lower box, the shaft was slightly unstable. This issue, coupled with a strong servo caused the filter to oscillate when it was moved in front of the camera. The microcontroller had to be altered in order to compensate for this instability and the shaft was further loaded to dampen any oscillations.

Future Suggestions

There are better ways to operate a digital camera than pushing buttons with servos. Using servos limits the number of buttons on the camera that can be pressed, uses a lot of power, and button pressing using a servo is not incredibly precise. It would be a more elegant solution to activate the buttons through some sort of electrical means. This would probably require opening the camera up and soldering wires directly to camera components. There is a risk of damaging the camera in this process, which is why it was not attempted this year.

A more efficient power system should be researched. Linear regulators were used to supply power to all of the electronics. These regulators get the job done, but are not nearly as efficient as some sort of switching power supply. During our flight test, the balloon was in operation and under full power for about one hour and it is estimated that the balloon could operate for 2 to 3 hours with the current power supply.

Image Processing

Background Information

The research goal of the Balloon Sat project is to use image processing techniques to extract data about some particular ground feature on Cal Poly's campus. The Balloon Sat project has been attempted by several teams of students in past years, each choosing to focus on a different aspect of feature extraction and classification. The 2007 Balloon Sat team has sought to launch the balloon to an altitude of about 1000 feet and proceed to take a series of photographs.

A photograph is taken when the camera's shutter opens allowing the light from a scene to enter the camera and affect the film. In digital camera, the "film" is actually a device called a CCD array ("Charge Coupled Device"). Photons (light waves) reflected or emitted from the scene enter the camera, and each photon has a certain amount of energy. When the photon hits the CCD, with enough energy it may excite an electron which flows into a pixel well within the CCD array. This process continues while the shutter is open, so that each well begins to fill with electrons. When the camera shutter closes, the CCD array releases to memory the percentage filled for each well. When the well is almost completely filled, the camera stores a bright pixel. Conversely, when the well is empty, the camera stores a dark pixel. The CCD contains pixel wells for every pixel in the image. So, for our 10 Megapixel camera, the CCD array contains 30 million wells (each pixel must have 3 wells for red, green, blue color)!

Before the camera can actually receive data, the scene must be illuminated by a source (in our case, the sun). Then, each object in the scene reflects certain colors. For example, a blue car reflects blue light, so blue photons enter the camera and excite the CCD. The end result is a number of blue pixels stored in the digital image. The human eye uses color response to identify objects in a quite similar fashion. The range of wavelengths reflected by a particular object is called its spectral response. It is important to note that most objects reflect light outside of the visible spectrum. However, the human eye is physiologically limited to what it can see.

In general, digital cameras are responsive to a wider spectrum of light than humans. The human eye can detect light in the wavelength range of 400 nanometers (blue) to 700 nanometers (red), but a camera can detect light in the infrared range (greater than 700 nm). Therefore, if we could block visible light entering the camera, the camera could take pictures recording light exclusively in the infrared region. Since a traditional image provides only red, green and blue colors for object identification, we would traditionally be limited to using those three "spectral bands." However, by taking pictures exclusively in the infrared region, we would gain a dimension in analysis which could greatly benefit the accuracy of feature extraction. Unfortunately, most cameras come equipped with an internal infrared-blocking filter, so taking reliable pictures in the infrared band requires longer exposure time, which will likely degrade image quality.

"Feature Extraction" refers to the broad scope of techniques used to identify components of an image of particular interest. Feature extraction often involves pixel manipulation, statistical analysis, and filtering. It is through feature extraction that we hope to extract useful data on ground features at Cal Poly. Our primary project goal has been to take aerial photographs of campus parking lots at different times and develop software programs to automatically calculate

the parking capacity and percentage full throughout the day. The university could use this data to plan for future parking needs, and students could plan when to drive to campus and where they are most likely to find parking.

THEORY

Our research has focused on developing reliable methods of feature extraction for identifying ground features. This section will explain some of our present analysis of general parking lot features and detail our methods of feature extraction.

Parking Lots

Much of image processing and machine vision theory has focused on trying to develop techniques that recognize image features in the same way as the human eye. The difficulty lies in the fact that our eyes have been training to recognize objects since our earliest cognitive experiences, and computers are asked to accomplish the same recognition task in a few seconds. To overcome this challenge, one must find a way to translate human perception into software. Thus, when developing methods for image processing, researchers often begin by visually analyzing the scene of interest to find cues that may assist in feature extraction. Similarly, our analysis begins with the consideration of known parking lot features.

Dealing with cars in a parking lot is both beneficial on some levels and more challenge on others.

As an example, we know that for the most part:

- Parking spaces are usually the same size.
- Parking spaces are perfectly aligned.
- Parking spaces are evenly spaced along each row.
- Parking spaces are clearly marked with paint stripes.
- Parking lots are characterized by gray to dark gray asphalt.

However, while these features benefit the process, we also note the following:

- Parking spaces are both diagonal and perpendicular.
- The condition of asphalt in the parking lot varies.
- Stripes may be white, yellow, or hardly visible.
- Poorly parked cars often cover stripes.
- Rows are not necessarily the same length and are not necessarily parallel to one another.
- Lines in the parking lot will probably not be vertically aligned; this feature is wholly dependent on the camera's orientation with respect to the ground.

Additionally, each year consumers have more choices in choosing the body style and color of their cars. We must keep in mind all the different possibilities.

Methods of Feature Extraction

The following methods are an example of pattern recognition techniques currently in development. Each technique will likely provide some useful information in feature extraction. In the end, we hope to piece these together to maximize the accuracy of our analysis software.

Grayscale Conversion

The Canon G7 camera, like all modern recreational digital cameras, is preset to take color pictures. Often, however, feature extraction requires a grayscale image instead of full color. There are many ways to compute this conversion. Most algorithms involve using a linear combination of the red, green and blue pixel values to compute the gray value. In our case, we chose to borrow that RGB□YUV algorithm as defined by the JPEG standard to calculate the intensity (Y) part of the YUV color space and use Y as our gray value.

$$gray = 0.299red + 0.587green + 0.114blue$$

Edge Detection

Human perception relies heavily on the discernment of object edges. Similarly, many feature extraction techniques include edge detection as a step in the algorithm.

Convolution is a mathematically intensive operation involving many multiplies and adds. Edge detection is an example of a 2-D discrete convolution, where each local image region is convolved with a "mask," a feature extraction tool. This process is repeated for every pixel. Thus, a large mask can require a considerable amount of processing time.

Many sophisticated edge detection algorithms are available but unnecessary for our purposes. We chose a simple Laplacian mask as our edge detector:

-1	-1	-1
-1	8	-1
-1	-1	-1

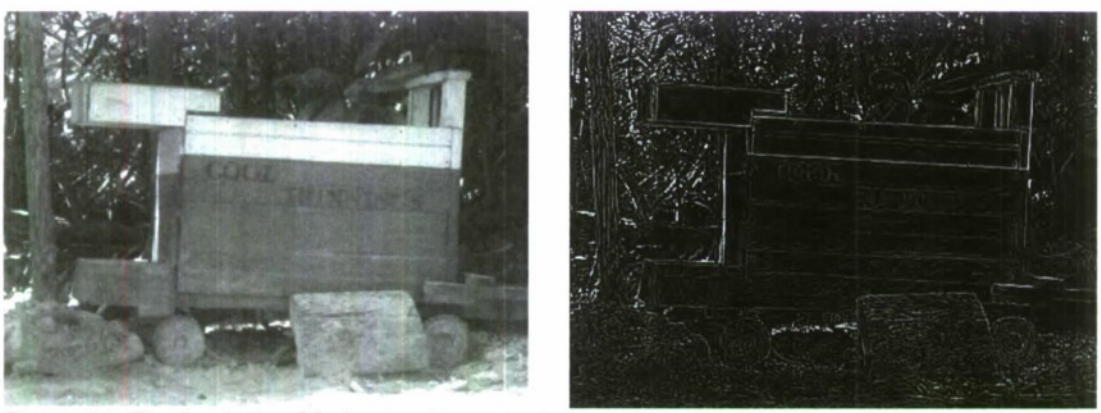


Figure 12: The Laplacian Mask edge detector enhances dominant edges in an image.

Thresholding

Sometimes it is beneficial to obtain a binary image (exclusively black and white) for the purpose of processing or counting. Thresholding converts any grayscale image to a binary image. All pixels can be described by a brightness value in the inclusive integer range of 0 to 255, where 255 corresponds to the brightest pixel (white) and 0 corresponds to the darkest level (black).

To threshold, simple choose a threshold level between 0 and 255 (200, for example). Any pixel value greater than 200 will now be set to 255 (white), and any pixel darker than 200 will be set to 0 (black).

Figure 13: A threshold level of 170 brings brighter image regions to the image foreground.

Hough Transform

The Hough Transform process is used for identifying unknown straight lines (and sometimes other shapes) in an image. We used it for finding the parking lot orientation with respect to the camera.

The Hough Transform requires a binary (thresholded) image of the edge features in the original image. It takes every white pixel with coordinates *(x,y)* in the binary image and translates that pixel into a sinusoidal wave in Hough space [coordinates *(θ, ρ)*], using the equation $\rho = x\cos\theta + y\sin\theta$.

Most people are familiar with the $y = mx + b$ notation for the equation of a line, but as lines tend toward vertical, the slope *m* goes to infinity. Therefore, the Hough transform makes use of the normal representation of a line. In normal representation, a segment is drawn from the origin to the line and is perpendicular to the line. The length of that segment is called ρ. θ, then is defined as the angle to that segment from the x-axis (positive angles are defined as counter-clockwise.

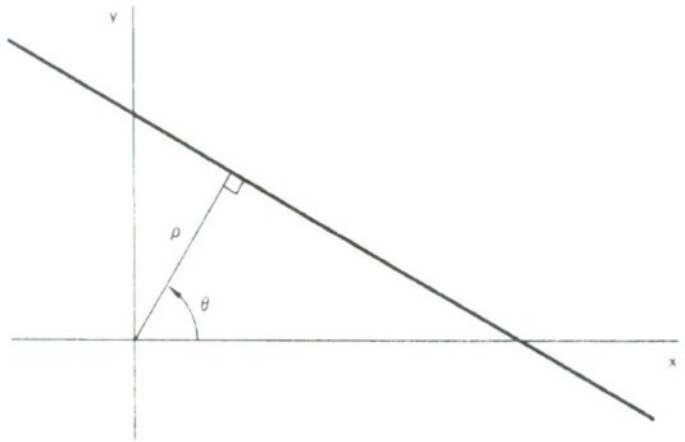


Figure 14: In the x-y plane, ρ is the normal (perpendicular) distance to a line of interest.

In an image, all pixels reside in quadrant 1, so that x and y are always positive. Thus, all possible lines in the image can be described by angles in the range $0^\circ \leq \theta < 180^\circ$. ρ can be positive or negative, but the image dimensions determine its limits.

Hough space represents a 2-dimensional image graph of all white pixels in the image. Each Hough space is divided into small regions called accumulator bins. The bin size is defined by the user, and all bins have an initial value of zero. As each sine wave is plotted, whenever the wave lands in a Hough space bin, the value of that accumulator bin is incremented by one.

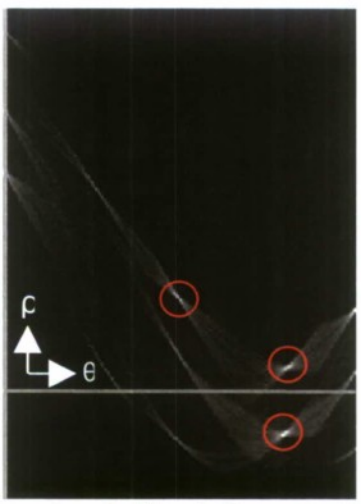


Figure 15: Bright regions in Hough Space are of primary interest.

If the image contains lines, then certain accumulator bins will have significantly larger values than others. When the Hough space image is finally generated, bins with the largest value are given the brightest pixels and correspond to the most probable lines in the image, circled in red in the example.

At this point, the savvy observer would reeognize that parking space lines are parallel. That is, spaees have the same angle orientation. Considering the normal representation of a line, this means that although ρ will not be the same for each parking stripe, θ should remain consistent. In Hough space, we would expect to find bright points oriented vertieally.

For our purposes, knowing the normal distance to the image lines does not seem too useful. However, the orientation angle assoeiated with those bright points in Hough spaee is a vital part of parking lot reeognition. So, the Balloon Sat method takes the top 50 accumulator bins and returns the angle assoeiated with greatest number of high-valued bins. In tests, this angle has consistently proved to be the true parking lot orientation.

Rotation

One the orientation angle is known, a rotation algorithm must be implemented so that the final produet is a vertically aligned parking lot. Rotation is primarily a two-stage process.

First, all digital images have reetangular dimensions corresponding to the horizontal and vertical size. If one rotates the image, the rectangular dimensions will no longer eorrespond to the horizontal and vertieal size. As a result, the first step involves the ealeulation of the new image dimensions. Since the total area will be greater than the initial area, some regions in the new image will inevitably contain no image data.

Then, once the new image is created, at each pixel in the rotated image, a eoordinate transformation finds the pixel location in the original image. In the following transformation matrix equation, *(x',y')* are the transformed eoordinates, *(x,y)* are the original coordinates, and θ is the rotation angle (counterelockwise angles are positive).

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

Sinee cosine and sine return real, non-integer values, the ealeulated x' and y' are not likely to be located right on an integer pixel. Thus, a bilinear interpolation of the pixel neighborhood is used to approximate the aetual output pixel value.

Correlation

Perhaps the most forefront teehnique developed thus far has been correlation. A eorrelation ealeulates how alike two objeets are. If we develop a "key", some partieular feature of interest, and run a correlation over the entire image, we will find how often it is present in the image. Suppose that feature is a general ear. We choose a prototype (a best ear in the seene) that looks most like all the other ears. After running the correlation, we hopefully would find all the places throughout the seene where a ear is present.

Correlation, like edge deteetion, involves a 2-D eonvolution. The mask, in this ease, is just the prototype, but the eorrelation must still be ealeulated for every pixel location in the output image. Thus, large prototypes may require notieeably large proeessing times. The eorrelation method can only work if images have been vertieally aligned, which stresses the importance of the Hough Transform and rotation methods diseussed earlier.

In essence, the correlation result is measure of how much the image region and the prototype are similar. If the comparison is strong, the correlation produces a bright region. If the prototype and image seem to be opposite in color over the region, the correlation returns a dark region. If there seems to be no comparison between the prototype and image region, the correlation returns a bland gray region. To the untrained eye, a correlation image will often appear as a blurry grayscale version of the original.

To extract data from a correlation image, a threshold must be applied. This step suppresses regions of poor correlation and highlights regions of interest. Ideally, the resulting image is full of small bright regions that represent each car matching the prototype. These may now be counted, and that number should equal the number of similar cars in the image.

Statistical Analysis of Parking Spaces

Suppose through feature extraction we identified the location of a parking space. We believe we can determine whether or not a car sits in the parking space through a simple statistical analysis. This method capitalizes on the fact that most parking spaces are gray asphalt. By observing the mean and standard deviation of the pixels in that parking space, we can identify the most prominent color (the mean) and whether or not the colors present are spread over a wide range.

This feature extraction technique was not explored by the 2007 Balloon Sat team, because techniques described previously exhausted time resources. The correlation method is insufficient for finding parking spaces in an image, and they must be found before they can be statistically analyzed. We hope that this avenue is explored more fully next year.

Infrared Signature

Many people associate the term infrared with heat emission. While the true spectral signature of heat lies primarily above the infrared response of the camera, we may be able to identify "warm" vehicles (those recently parked) through thresholding infrared images. This method has not yet been tested.

Additionally, applying the infrared image to feature extraction techniques could add an entire dimension of available information. Some features not characteristic of the blue through red wavelengths might appear in the infrared band region.

Results

What follows is a description of the image processing algorithm developed by the 2007 BalloonSat team. Please note that in the absence of test data, the software was developed using an aerial image obtained from the internet.

Figure 16: This parking lot has been chosen for processing.

The BalloonSat Algorithm

Step 1: Grayscale Conversion

Before correcting the alignment, the Hough algorithm requires grayscale images without degradation in image quality. The JPEG standard conversion specifies that RGB values are converted to gray level values using $gray = 0.299red + 0.587green + 0.114blue$.

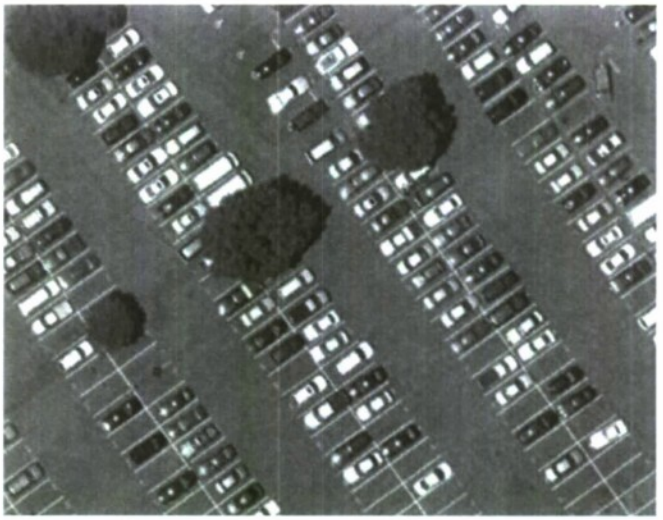


Figure 17: Grayscale conversion has no resulted in a loss of image quality.

Step 2: Edge Detection

Finding parking lot orientation requires the enhancement of edge features, regions in the image that correspond to rapid changes in the gray level brightness.

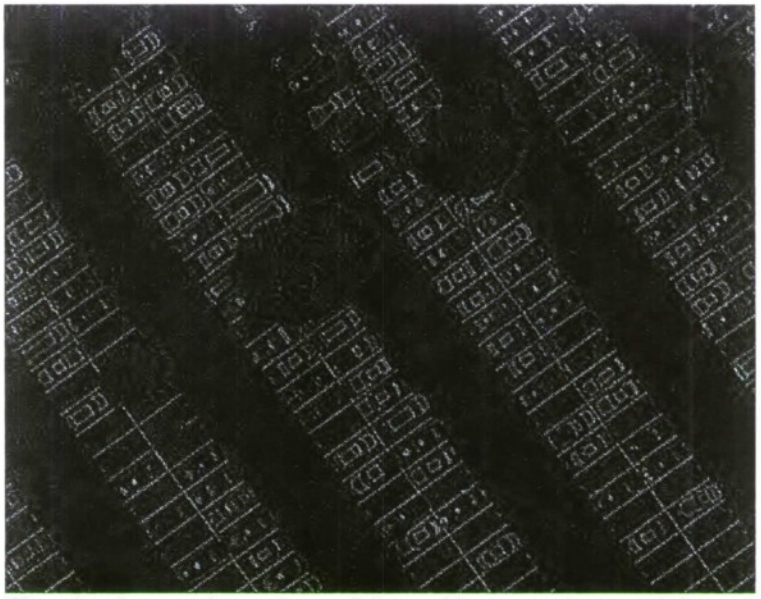


Figure 18: The edge detector has enhanced the parking space lines and the outlines of cars.

Step 3: Threshold

Though it may be difficult to see, the edge image contains various gray levels. The Hough Transform requires a simple binary decision for its translation into Hough Space. A threshold converts many of the bright pixels to pure white and forces darker pixels into the background.

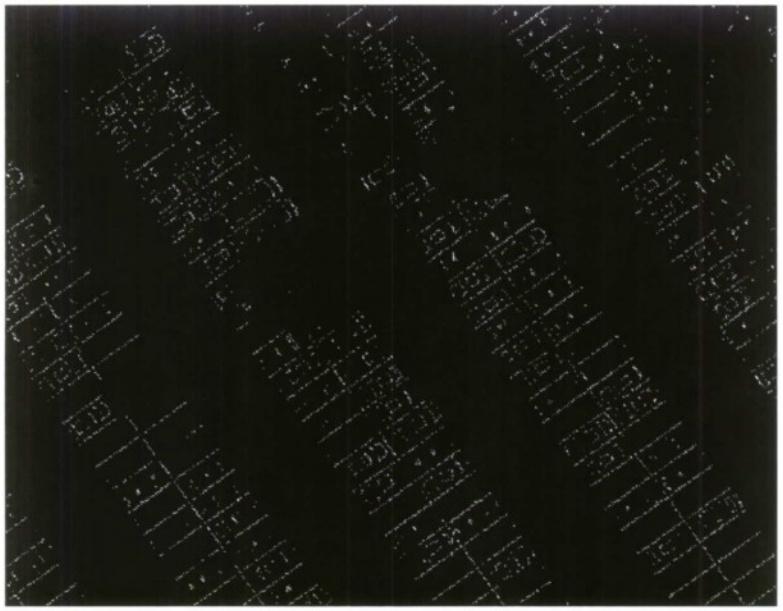


Figure 19: After a threshold of 200, the parking lot is ready for Hough processing.

Step 4: Hough Transform

To find dominant angles, all white pixels with coordinates *(x,y)* are translated into Hough Space using the equation $\rho = x\cos\theta + y\sin\theta$. The Hough Space graph shows probably lines as bright points, and in the case of parking lots, we expect these bright points to line up vertically, as all parking spaces are painted at the same angle – parallel to one another. The result of Hough processing is an orientation angle, which will then be used for alignment.

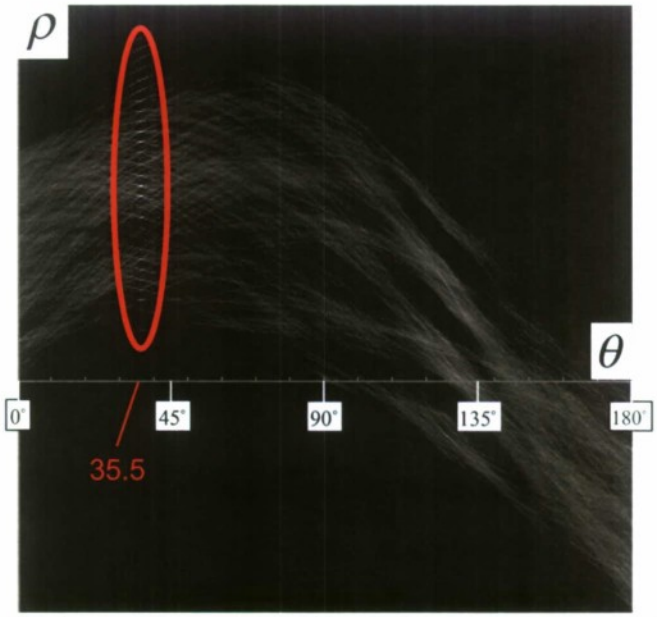


Figure 20: Vertically aligned bright points in Hough Space reveal the parking lot orientation.

Step 5: Rotation and Alignment

With the knowledge of the orientation angle, one may now correct the image alignment. The parking lot is rotated to the best angle, and outlying regions are set to gray.

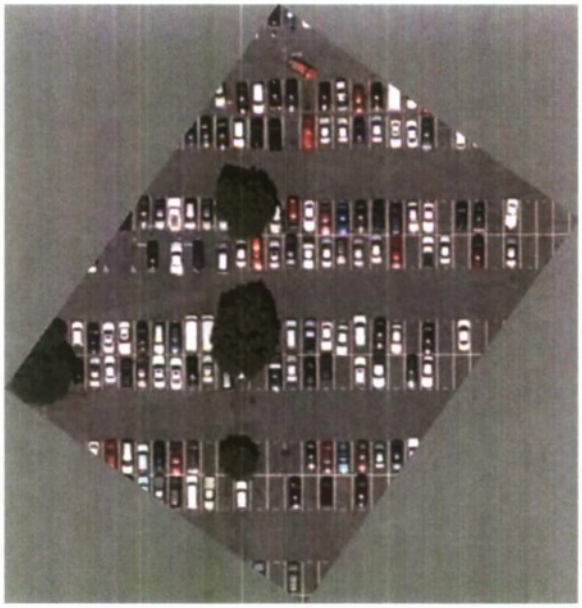


Figure 21: Rotation using the angle from Hough processing results in a vertical alignment.

Step 6: Correlation

Using the vertically aligned parking lot and a chosen prototype, the correlation image is calculated. Bright regions are the primary regions of interest after correlation. Though the image resembles a blurry version of the original, the information contained in a correlation is quite different. Bright regions represent regions of strong comparison between the region and the prototype.

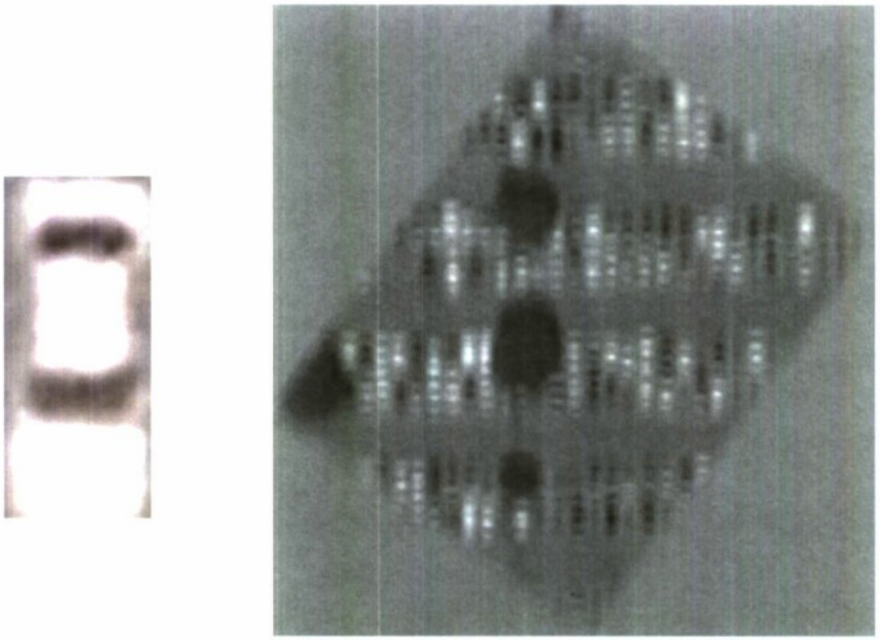


Figure 22: Correlation with the white car prototype enhances regions of strong comparison.

Step 7: Thresholding

To count the cars in this image, another threshold step is necessary. A threshold isolates the regions of interest and ideally these regions represent cars that agree with the prototype.

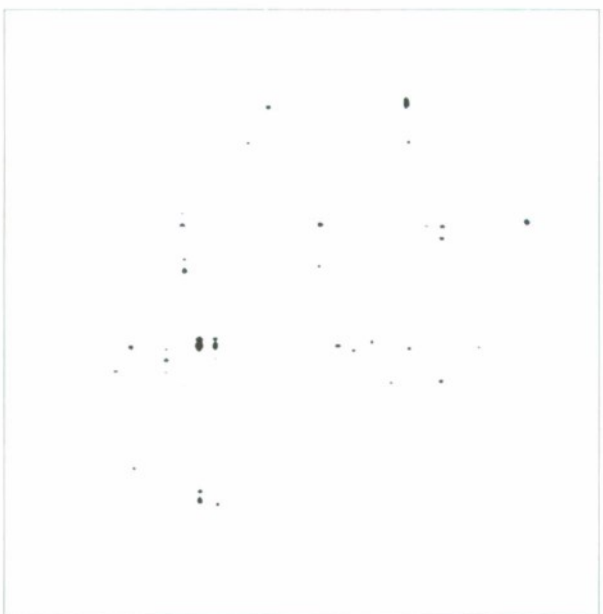


Figure 23: Thresholding the correlation image reveals probable locations of white cars.

Step 8: Counting

A counting algorithm identifies isolated regions in the thresholded correlation. This number should equal the number of prototype-like ears in the original seene.

Figure 24: The analysis software identifies 36 regions corresponding to probable car locations.

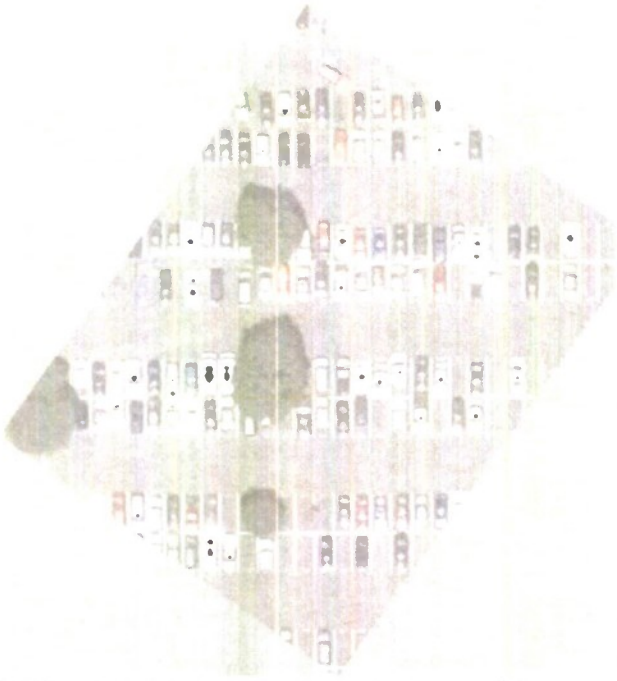


Figure 25: The recognition overlay shows that many white cars were identified, but some were missed and some were counted twice.

Shortcomings

As seen in the reeognition overlay above, the white prototype failed to identify all the white cars in the image. Furthermore, some ears were identified twiee, which is the ease when two black regions are loeated on a ear in the reeognition overlay. We also notieed that the correlation works best when trying to locate white ears. Other ear eolors seem elose enough to the colors of foliage that trees pass through the threshold. The eorrelation method is therefore insufficient for the kind of results we need.

We found that the Hough Transform algorithm is also limited in its usefulness due to the needs of the edge detector. For effective edge detection using the Laplacian mask, the image must have high contrast between the background (asphalt) and the paint stripes. Unfortunately, for the camera to collect enough light data, the CCD ISO response must be set high. This often results in a background image washout, in which case the asphalt approaches the color white – nearly the same color as the paint stripes.

When images like these pass through the edge filter, the enhanced edges can still be found in the image but are extremely dim. Correcting these dim edge images would certainly require histogram processing, but optimizing camera settings may help to avert the situation altogether.

Future Improvements

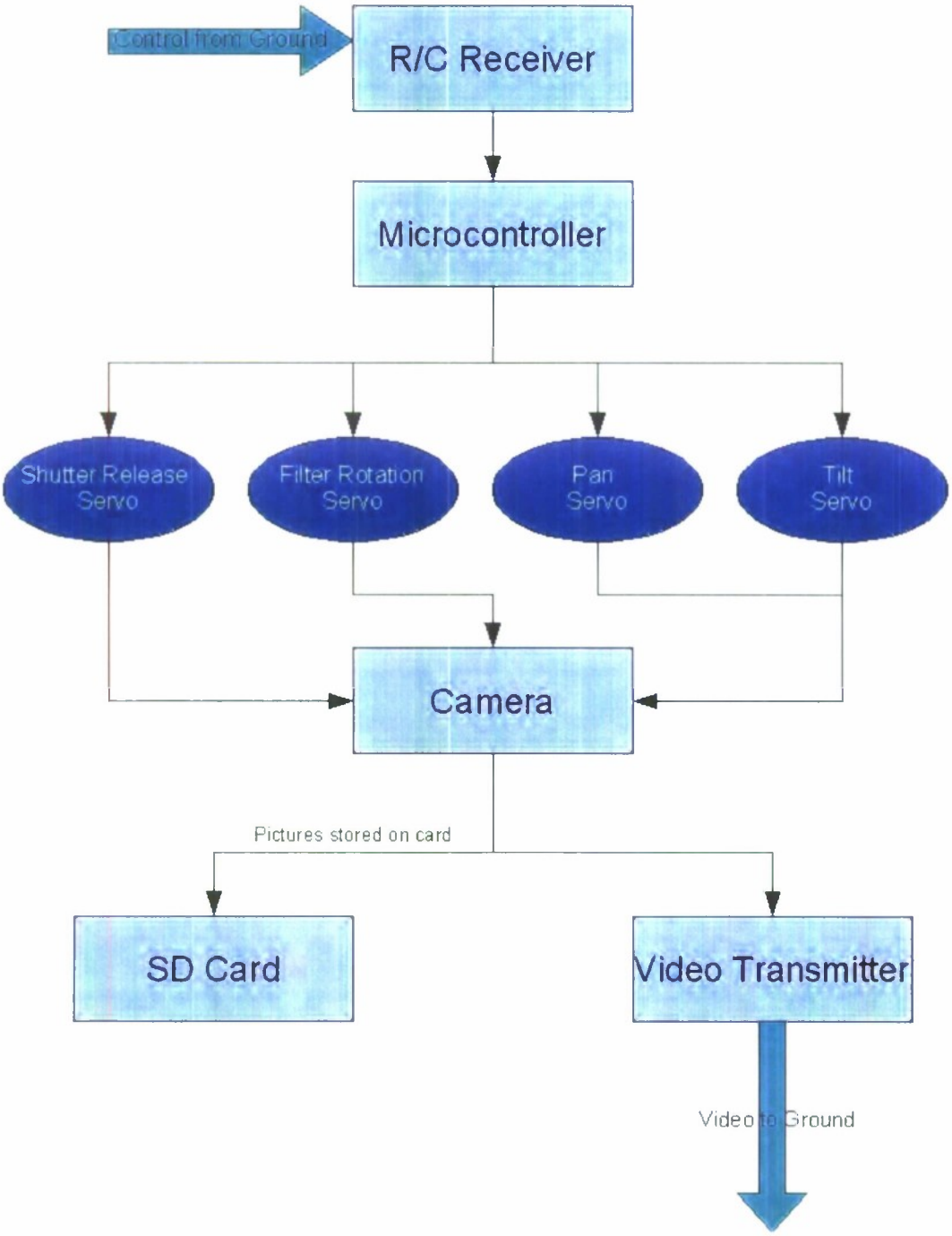
Next Year's BalloonSat team should consider improvements to the traditional correlation algorithm, which favors extreme pixel values over mid-tones. Without drastic improvement, the correlation will continue to fail in identifying the majority of cars in a parking lot.

Difference imaging would be a great step to take in the future. The theory is relatively simple. Say one has an empty parking lot image and also an image of the same parking lot with a few cars. In each region, compare the empty lot image with the populated lot image. If pixel differences are large, output a brighter pixel. If differences are dim, output a dim pixel. This idea would require some intense image registration – which could possibly be handled by some sort of correlation – because the differencing only works when the empty lot and populated lot images are perfectly aligned…rotationally, vertically, horizontally, and in pixel ground cell size.

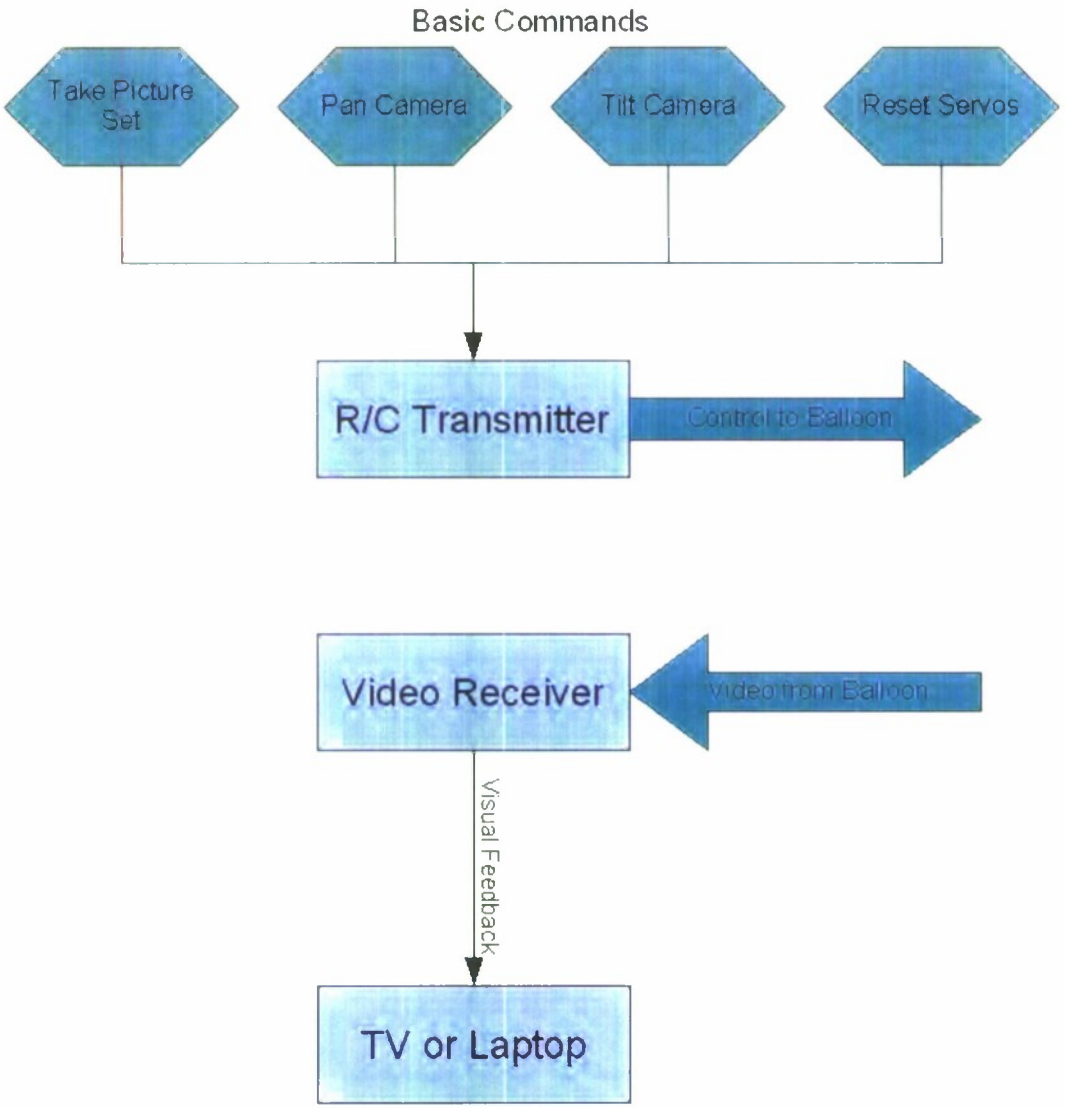
To count parking spaces, an entirely new algorithm must be developed. Once a parking space is identified and located, statistical analysis should prove quite effective in determining whether a car is located in the parking space or not.

Appendix A – Block Diagrams

Onboard Block Diagram



Ground Station Block Diagram



Appendix B – Electronics Parts List

Electronics Parts List *					
Purpose	*Component(s)*	*Jameco part#*	*Qty.*	*Unit Price*	*Price*
	Onboard Electronics				
R/C Receiver	GWR-6NII/H 6CH 75MHz R/C FM Receiver	N/A	1	$21.22	$21.22
Video TX/RX	1.2Ghz Composite Video Transmitter w/3000ft range	N/A	1	$57.99	$57.99
Batteries	NiMH or Lithium Rechargeable	N/A	1	$34.88	$34.88
	Battery Boxes	N/A	1	$4.07	$4.07
Servos	SG5010 digital hobby servos	N/A	2	$9.01	$18.01
	SG90 mini hobby servos	N/A	2	$6.49	$12.98
Microcontroller	Sample: PIC18F1320	N/A	1	$0.00	$0.00
	18 pin IC socket	38113	10	$0.10	$0.99
PC Board	Donated: Double sided 1oz copper clad FR4 Fiberglass	N/A	1	$0.00	$0.00
Resonator	20MHz ceramic resonator w/built in capacitors	324671	2	$0.32	$0.64
5V Voltage Regulator	7805 1A Linear Regulator	51262	2	$0.35	$0.70
12V Voltage Regulator	PQ12RA11 1A ON/OFF Regulator	Digi: PQ12RA11	2	$1.95	$3.90
Plug for Video Xmitter	2.1x5.5mm Female Power Plug	191484	1	$0.69	$0.69
Diodes	1N4001 1A Rectifier	N/A	3	$0.00	$0.00
Capacitors	47uF	Digi: P13458-ND	3	$0.11	$0.33
	Donated: 0.1uF	N/A	2	$0.00	$0.00
Resistors	Donated: 47kohm	N/A	1	$0.00	$0.00
Receptacle for Receiver	.100" (2.5mm) Female Header Receptacle	308567	2	$0.69	$1.38
Headers for Servos	.100" Straight Male Headers (Gold)	53532	2	$0.59	$1.18
ON/OFF Switch	DPDT Slide Switch	621481	1	$1.65	$1.65
Shipping for Electronics	Shipping and Handling	N/A	1	$14.81	$14.81
	Ground Station Electronics				
R/C Transmitter	Borrowed: VEX 6CH 75MHz FM Transmitter	N/A	1	N/A	$0.00
Monitor	TV or Laptop	N/A	1	N/A	$0.00

Total Cost
$175.42

Appendix C – MikroBasic Code

```
program BalloonSat_Firmware_V2

dim    i,
       j as byte
     in0,
     in1,
     in2,
     in3,
     in4,
     in5,
  center as integer
    serv as integer[4]
   servo as integer[4]

const PRESS = 50
const FOCUS_TIME = 100
const TRANS_TIME = 100
const PICTURE_TIME = 50
const FOCUS = 45
const PICTURE = 25
const RES = 2
const DELAY = 5
const FILTERS = 1
const INC_FILT = 160

sub function move_pantilt(dim in as integer) as integer
    if in < (center - PRESS) then
         result = (in - center) >> 5
    else
        if in > (center + PRESS) then
             result = (in - center) >> 5
        else result = 0
        end if
    end if
end sub

sub procedure read_receiver

    in0 = 0          'clear input values
    in1 = 0
    in2 = 0
    in3 = 0
    in4 = 0
    in5 = 0

    while TestBit(PORTB, 3) <> 1       ' wait for pulse on CH1
    wend
    while TestBit(PORTB, 3) = 1        ' wait until CH1 goes low again
          Delay_us(RES)
          in0 = in0 + RES               ' measure length of CH1 pulse
    wend
    Delay_us(1)
    while TestBit(PORTB, 2) = 1
          Delay_us(RES)
```

```
            in1 = in1 + RES                ' measure length of CH2 pulse
      wend
      while TestBit(PORTB, 7) = 1
            Delay_us(RES)
            in2 = in2 + RES                ' measure length of CH3 pulse
      wend
       Delay_us(1)
      while TestBit(PORTB, 6) = 1
            Delay_us(RES)
            in3 = in3 + RES                ' measure length of CH4 pulse
      wend
      Delay_us(1)
      while TestBit(PORTB, 5) = 1
            Delay_us(RES)
            in4 = in4 + RES                ' measure length of CH5 pulse
      wend
      Delay_us(1)
      while TestBit(PORTB, 4) = 1
            Delay_us(RES)
            in5 = in5 + RES                ' measure length of CH6 pulse
      wend

end sub

sub procedure output_servos
      serv[0] = servo[0]
      serv[1] = servo[1]
      serv[2] = servo[2]
      serv[3] = servo[3]
      PORTA = %00000001
      while serv[0] > 0
            Delay_us(RES)
            serv[0] = serv[0] - RES
      wend
      PORTA = %00000010
      while serv[1] > 0
            Delay_us(RES)
            serv[1] = serv[1] - RES
      wend
      PORTA = %00000100
      while serv[2] > 0
            Delay_us(RES)
            serv[2] = serv[2] - RES
      wend
      PORTA = %00001000
      while serv[3] > 0
            Delay_us(RES)
            serv[3] = serv[3] - RES
      wend
      PORTA = %00000000
end sub


main:
      Delay_ms(2000)

      ADCON1 = $FF                     ' PORTB all digital
```

```
    TRISB = %11111100              ' configure RB2-RB7 as inputs
    TRISA = %11110000              ' configure PORTA pins as outputs
    PORTA = %00000000              ' initialize servos
    PORTB = %00000010              ' turn video transmitter on

'-----------------------------------
    'FIND CENTER                    test for a center input pulse
    in0 = 0
    while TestBit(PORTB, 3) <> 1   ' Ensure that transmitter is on
    wend
    while TestBit(PORTB, 3) = 1    ' Ensure that transmitter is on
    wend
    Delay_ms(500)
    while TestBit(PORTB, 3) <> 1   ' wait for pulse on CH1
    wend
    while TestBit(PORTB, 3) = 1
        Delay_us(RES)
        in0 = in0 + RES            ' measure length of CH1 pulse
    wend
    center = in0
'-------------------------------------

    servo[0] = center - (center >> 3)
    servo[1] = center - (center >> 3)
    servo[2] = center - (center >> 3)
    servo[3] = center - (center >> 3)

here:
    read_receiver

    servo[0] = servo[0] + move_pantilt(in2)
    servo[1] = servo[1] + move_pantilt(in3)

    if in4 > (center + PRESS) then
       PORTB = PORTB xor %00000010
       DELAY_ms(200)
    end if

    if (in4 < center - PRESS) then    ' reset servos?
       servo[0] = center - (center >> 3)
       servo[1] = center - (center >> 3)
       servo[2] = center - (center >> 3)
       servo[3] = center - (center >> 3)
    end if

    output_servos

'-------------------------------------------
 ' PICTURE CYCLE
    if in5 > (center + PRESS) then     ' take picture?
       for j = 0 to FILTERS
           servo[2] = servo[2] + FOCUS     ' move servo to focus camera
           for i = 0 to FOCUS_TIME              ' wait to focus
               output_servos
               delay_ms(DELAY)
           next i
           servo[2] = servo[2] + PICTURE        ' move servo to take picture
```

```
            for i = 0 to PICTURE_TIME                ' wait to take picture
                output_servos
                delay_ms(DELAY)
            next i
            if j <> FILTERS then
                servo[3] = servo[3] + INC_FILT
                output_servos
            end if
            servo[2] = center - (center >> 3)
            output_servos
            DELAY_ms(TRANS_TIME)
        next j
        DELAY_ms(50)
        servo[3] = center - (center >> 3)

    end if
'-----------------------------------------------

    delay_ms(DELAY)

    goto here
end.
```

HONORS PROJECT

Design and Fabrication of a Resistance-Switching Memory Cell Based on Chalcogenide Glass

Project Investigators:

N.A. Vickers, W.P. Dunn, M.A. Jain and C. De Leo
(Richard Savage-Advisor)
Materials Engineering and Electrical Engineering
California Polytechnic State University
San Luis Obispo, CA

Executive Summary

There is a need for fast, non-volatile memory that can be scalable in order to create densely packed memory chips for high volume memory. Current flash RAM technologies cannot meet these requirments so research is going into new materials that look like they will be able to meet these needs. Chalcogenide glass is one such material that might meet these needs that has had much research go into it. The problem with this current technology is the lack of information that is needed inorder to integrate this new memory device with silicon microelectronic technology. Research done here at Cal Poly is focusing on looking at these issues in order to create a working device ready for market production. Research will progress in three phases. The first phase is the deposition and characterization of the chalcogenide thin films. The second phase is to manufacture the device and also test its electronic properties. The third phase is to integrate working device structures into a microfabrication process for market prototyping. Currently, the research is in phase two. Investigations in thermal anealing, oblique deposition and photocondensation have been done. Oblique deposition affects the amount of elemental sulfur present in the film. Photocondensation has no affect on the amount of elemental sulfur present in the film where annealing allows the sulfur to diffuse into the bulk material. Initial testing of the devices have been done and that they do exhibit the two phase electrical conduction. Further research will go into the affect of cross sectional area on device functionality, a further investigation of annealing on device functionality, and how thickness affects device fuctionality.

Introduction

. In the age of fast computers, the internet, and the growing need of digital memory for handheld devices such as digital cameras, and mp3 players, there is a need for ultra-dense non-volatile digital memory. The main candidate to solve this problem is flash memory, which comes in all shapes and sizes including our handy thumb drives, and memory cards for digital cameras. One limitation of flash memory is that it cannot be scaled like other microelectronic devices. At smaller dimensions, flash memory starts malfunctioning. This means that flash memory will not be able to meet high-density needs of the future. To solve this problem, a new type of technology is currently being research at universities and in industry.

Programmable metalization cell (PMC) memory also called ovonic memory or phase change memory (PCM) is that new type of technology. It utilizes new types of materials that can change its structure in order to store digital information. The material being research at California Polytechnie State University San Luis Obispo is silver doped germanium sulfide glass. The structure of this glass gives it unique electrical properties, which ean be used to store digital information. The purpose of this research is to create a low power non-volatile memory device that can be ultra scalable, and can function as well or better than eurrent flash RAM technology.

The subject of this report is to communicate the work currently being done towards the design and fabrication of a resistance-switching memory cell based on chaleogenide glass. The researeh is separated into three phases. The first phase is to deposit thin films of silver, aluminum, and silver doped chalcogenide glass, and characterize these films. The second phase is to fabricate a memory eell based on chaleogenide glass and eharaeterize its eleetrieal properties. The final phase is to investigate other materials with similar eleetronic behavior. The ultimate goal of this research is to create a design using the materials researched at Cal Poly for manufacture of these memory deviees, that will be able to integrate with current microelectronic processing techniques.

Background

Sinee the invention of the transistor, mieroelectronics industry has been focusing on minturization and integration. Not until recently has the focus ehanged to new types materials for the use in mieroelectronies. This is due to the fact that elcetronics are approaching a new limit. We can only make devices so small, so to satisfy the demand for speed and power efficiency new materials are being tested and used. One of these materials is silver doped ehaleogenide glass.

Chalcogenide glass is a material with very interesting and useful properties. One sueh property is that many chaleogenide glasses have nonlinear refraetive indexes, which make it an ideal material for optieal applieations. Other chalcogenide glasses sueh as AgGeSe, AgGeS, and CuGeS, have a two phase resistance to an electrical current. What this means is that at a low voltage aeross the material, it has the property of an electrolyte which has high resistance. After a certain voltage is reached its electrical properties will dramatieally ehange to a low resistive state. Usually this is a resistanee drop by about 10^5-10^7 Ω. This voltage level is called the threshold voltage. As the voltage is backed off the material stays in this low resistive state until it reaehes a maximum bias voltage. At this voltage the material reverts to its original state [1]. This phenomenon is very unique and interesting, and has many applieation as eomputer memory or other applieations like a chaleogenide glass emitter heterojunction transistor. However the strueture of this material is what deteremines these unique properties. A better understanding of its structure will lead to better utilization of these properties.

The strueture of this material is very interesting and unique. The bulk material has two phases after the metal (either silver or eopper) is doped in to the bulk. The main phase is a germanium rieh phase that is dominated by eovalent bonding between Ge and Se or S giving struetures similar to ethane and polyethylene [2]. The second phase is dominated by ionic bonding between the metal (silver or copper) and the ehalcogenide

(selenium or sulfur). This phase happens in tiny sphere like regions. When a voltage is applied aeross the materials that is greater than or equal to the threshold voltage the strueture of the material changes to silver atoms eonneeting the eleetrodes together [3]. This is the reason for the resistanee drop. On the boundaries between these two phases a different hybrid strueture is formed. They are named Q units, and there are four different ones. The most commonly oeeurring Q-unit is Q_3, where a Ag ion will replaee a Ge atom causing an ionic bond with the sulfur atom. These structures are important because they help tell how dispersed the silver is in the bulk, and possibly the amount of spheres that oecur.

Many different designs for deviee strueture have been used. The most popular one is a cross. A conductive inert metal is deposited on a substrate, and then the chaleogenide layer and the anode are deposited on top of the inert eathode so that the lines are crossing each other at a 90 degree angle. This is the structure that we are currently using. Another deviee strueture is a planar method where the anode and cathode are deposited on the substrate, and then the ehalcogenide layer is then deposited bridging the gap in between the two eleetrodes [4]. The advantage of the first device is that it is really eompaet, easy to fabrieate, and replieates the strueture of the cells when they will be employed in industry. The benefit of the seeond device strueture is that it is easy to image in order to see ehange in miero-strueture and nano-strueture.

Many techniques are employed to see the strueture and composition of these materials. Common ways for imaging are seanning eleetron mieroseopy (SEM) and tunneling electron mieroseopy (TEM). Both of these processes image the surface giving you a picture of some sort, so you ean see different phases and order of atoms. Other tools are x-ray diffraetion (XRD), and x-ray fluoreseence (XRF). Also Raman scattering is a very good tool when using organie eompounds, and glasses like ehalcogenides. Raman spectroseopy measures the intensity of light inelastieally seattered. When a photon hits a ehemieal bond, that bond loses a eharaeteristie amount of energy eorresponding to the different vibrational modes of that bond. A detector measures the seattered light and then graphs the intensity of eaeh ehange in wavelength. To aide in ealculations the x axis is usually the ehange in wavelength. In figure 1. different raman plots are shown for a GeS system with a varying amount of Ag corresponding to the weight pereentage on the right hand side of the graph.

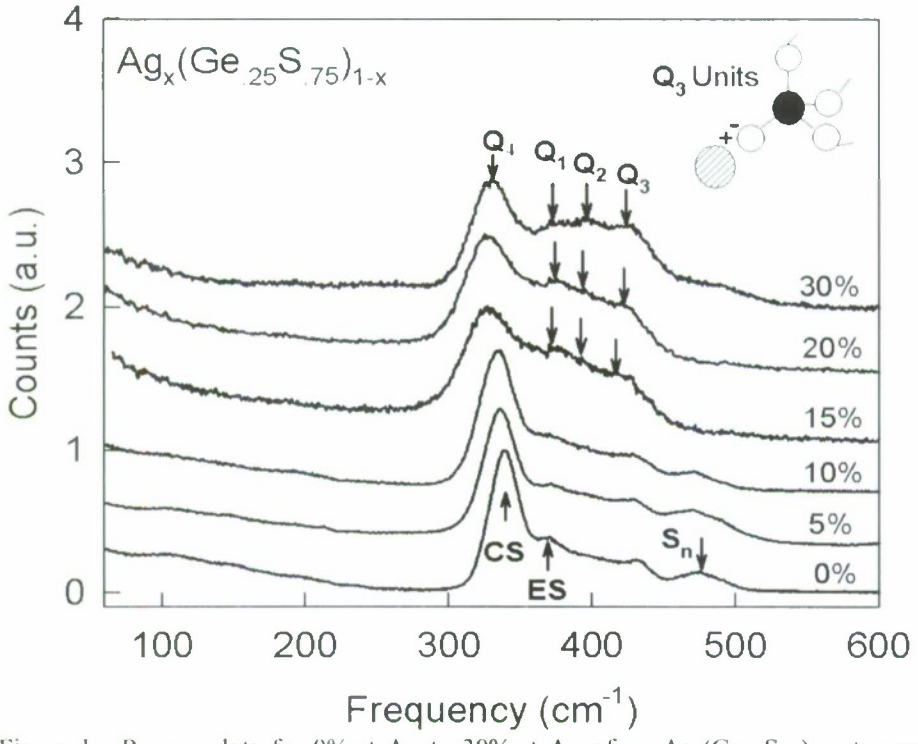


Figure 1. Raman plots for 0%wt Ag to 30%wt Ag of an $Ag_x(Ge_{.25}S_{.75})_{1-x}$ ternary system. The main distinguishing feature is the amount of Q_1, Q_2, and Q_3 units there are. The larger amount of silver correlates with a larger amount of these units.

Methods

The method currently used for depositing the films is vacuum thermal evaporator. A typical system is shown in figure 2. The crucible also shown in figure 2, holds the bulk material that we want to deposit, is resistively heated with tungsten mesh wrapped around it to keep bulk material from flying out. A beaker is used to hold the shadow mask and the glass slides which is the substrate that we are depositing on during the deposition.

To form the devices, we use a shadow mask as shown in figure 3. The areas that are cut out are where the film stays on the slide, elsewhere the film is lifted off with the mask. Also in figure 3 is a vile, that is used for storage, that contains some chalcogenide glass. A profilometer is used to measure the thickness of the films. For thermal annealing of the films, the samples are put on a hot plate for an hour at a certain temperature which is measured with a thermal couple. For photocondensation, the samples are put under a UV light source for an hour. All film thicknesses are measured using an Ambios XP-1 profilometer as shown in figure 4.

Testing of devices is done with the use of a probing station. The probing station is connected to a keithley meter which is a variable voltage source that measures current as well. Steps of .1 volts are done from 0V to 1V down to –1V and then back to 0V.

Figure 2. Left is a thermal vacuum evaporator. Used in depositing thin films. On the left is the crucible being held by two electrodes. Tungsten mesh is wrapped around the crucible to help keep the chalcogenide glass bulk material from flying out.

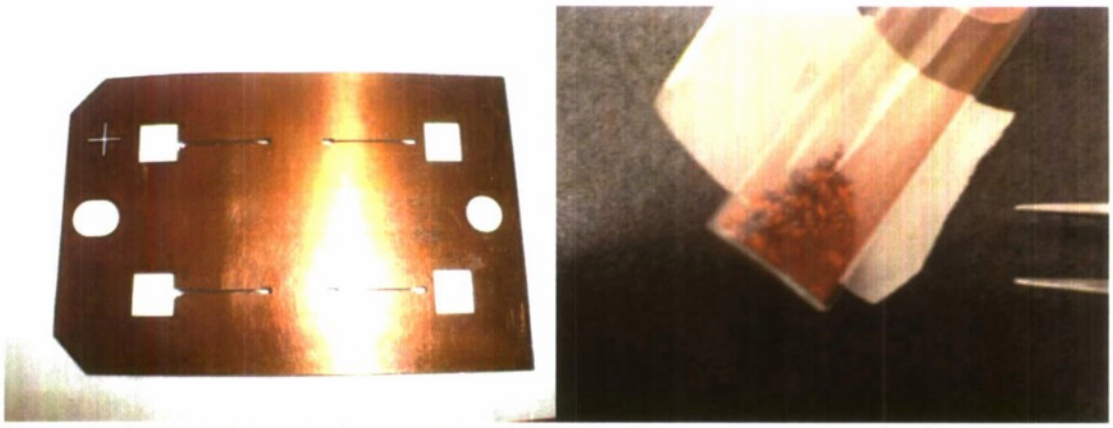


Figure 3. On the left is a shadow mask that defines the aluminum electrodes. On the right is a vile that contains some chalcogenide glass.

Figure 4. The profilometer that is used to measure thickness of the thin films.

Results

The thermal evaporator has been characterized so that the thickness of a film is known. For the chaleogenide glass it is almost a one to one ratio for the mass of bulk used and the thickness achieved in nanometers. During all of the depositions of the chalcogenide glass, we have had pieees of the bulk fly out of the erueible. This is easily explained by the lower evaporation temperature of the sulfur, so the sulfur is evaporating quickly which is ejeeting the bulk. Tungsten mesh is used to help prevent this from being a problem.

After a few depositions, we sent off some of the samples to have raman seattering done to analyze the ehemieal strueture of the films. Below are a eouple examples of the results that we received. The first two pietures are actual images taken of one of our films. The first photo is with a 10x magnifieation. The seeond image is the center of the first image but magnified 100x. the graph that follow the pietures, figure 6, shows the raman results for this part of the film. It is right on par with what we expected, however, there is much sulfur in our film eompared to the bulk. The sulfer can be seen by the sharp peaks at about 225 and 475. The other films that were deposited also had nice uniformity, and eonsistaney.

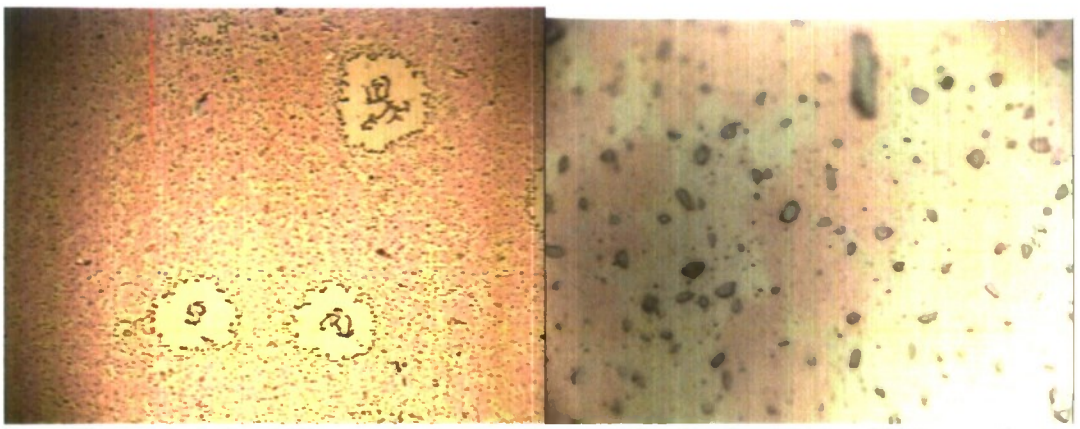


Figure 5. On the left is 10X magnification image of the film. On the right is a 100X magnification image of the film.

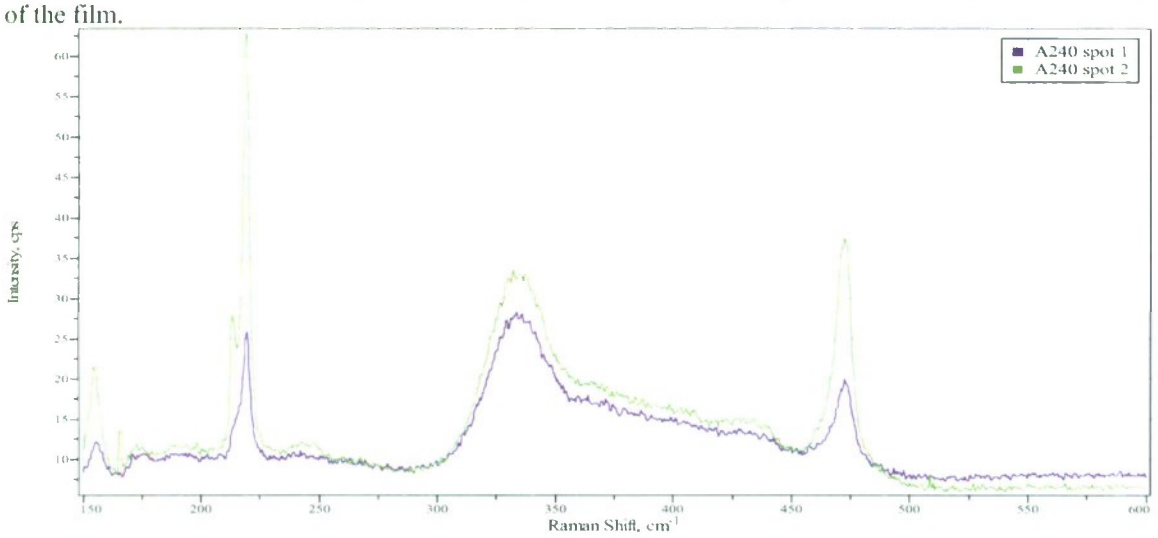


Figure 6. A graph showing the results of 2 different spots on the same slide. The large peaks on either side show the existence of S_8 while the curve in the center shows the existence of 10% silver AgGeS.

After finding out that the sulfur separates after deposition, different post depostition processes were tried. They are, photocondensation and annealing. It was found that photocondensation has no affect on the structure of the material. It was found that annealing does have affect. This affect can be seen easily through the raman data that the sulfur diffuses back into the bulk material. For a electronic device it is important for no elemental sulfur to be present as it will negatively affect the characteristics of our devices. Also we varied the angle of the substrate compared to the crucible. We found that sulfur is less prodominate at certain angles compared to others. The three angles that were tried were 90 degrees, 60 degrees, and 45 degrees.

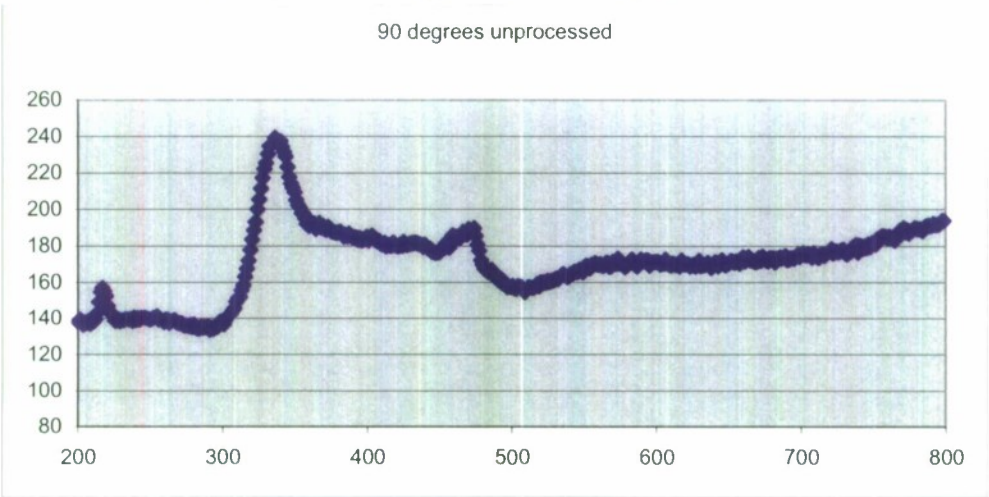


Figure 7. Results of raman spectroscopy done on a sample that was oriented at 90 degrees and had no post-deposition processing. As you can see there is a distinct sulfur peak at about 225 cm^{-1}.

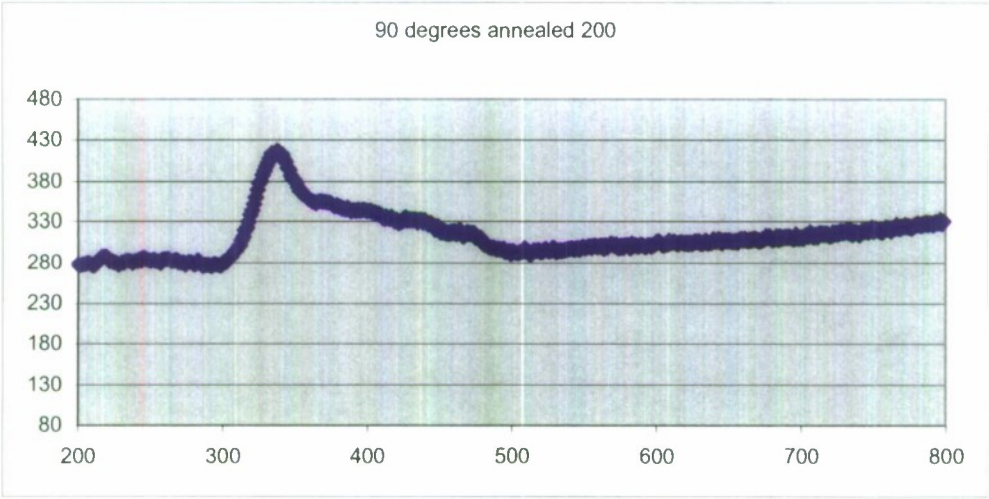


Figure 8. Results of raman spectroscopy showing no sulfur peak after the sample was annealed at 200 degrees fairenheit. It is infered that the sulfur diffused into the bulk because sulfur evaporates at 831 degrees fairenheit.

As can be seen from the previous graphs that the elemental sulfur goes away after annealing. Since a temperature of 831 degrees is needed to evaporate the sulfur which is 631 degrees above our processing temperature, it is believed that the sulfur diffused into the bulk material. This trend was observed for all orientations. The following chart summarizes the results of the thin film work.

Table I. Summary of Results for thin Film Characterization

angle	No proeess	photocondensation	150°F annealing	200°F annealing
90°	Sulfur present	No effeet	Sulfur diffuses partially	Sulfur diffuses completely
60°	Sulfur present	No effeet	Sulfur diffuses partially	Sulfur diffuses completely
45°	Sulfur present	No effeet	Sulfur diffuses partially	Sulfur diffuses completely

To understand more about the material and what eauses this two phase resistance phenomenon, our samples were imaged using scanning electron mieroscopy. Unfortunately the material phases are two small to eompletely resolve however useful information is still obtained. The following image is the results of doing SEM on the film. White fuzzy dots can be barely made out. Their approximate width is about 50 nanometers. These fuzzy white dots are the Ag_2S regions that are predieted. The reason they show up white on the SEM image is beeause they are the only eleetrieally eonduetive part of the film. Sinee the film is on a glass slide and that the rest of the film is a ehaleogenide glass, charge will build up where conduction can oeeur. The interesting information from this I that evaporation makes Ag_2S regions smaller and more dispersed through the bulk. Also the size of these regions in un evaporated material is on the seale of a mieron which is 200 times larger.

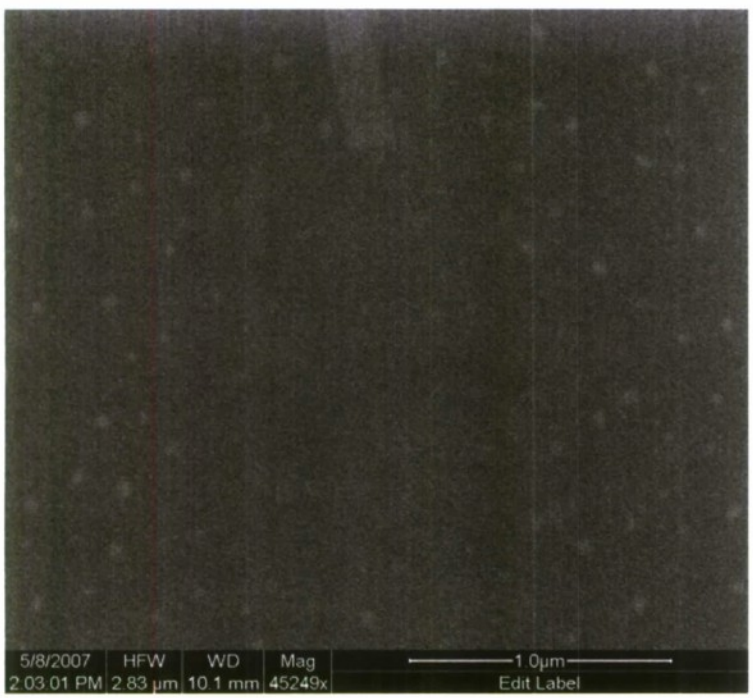


Figure 9. Scanning Electron Microscope image of the chalcogenide film. Fuzzy white dots are where charge build up occurs. In the case of this sample that is where the Ag_2S regions are.

Devices were successfully built and characterized. As was outlined in the background, these devices do exhibit a two phase resistance, however after a couple cycles they degrade and become non-functional. We also found that some thicknesses work better than others. Following is the results of one of our testing cycles. As can be seen is that there are two distinct phases with dramatically different resistance values. Also we found that thicknesses between 15nm to 30 nm have the best characteristics. Further research is going into these two thicknesses.

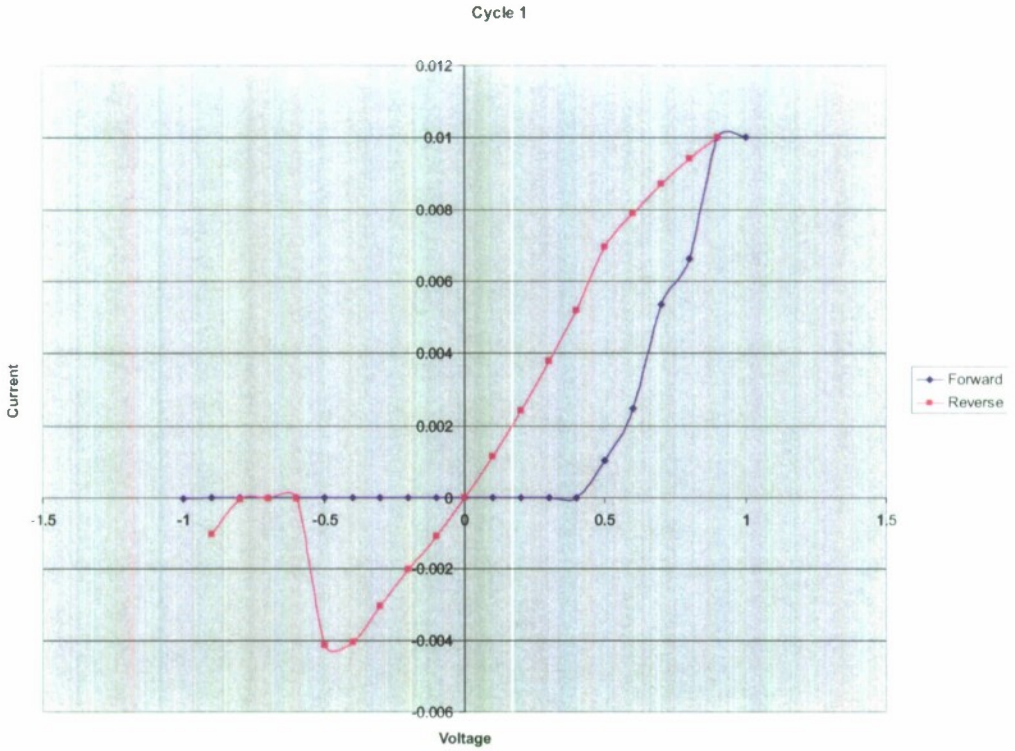


Figure 10. I-V relation for the resistance switching memory cell. As can be seen there are two different resistance levels. The high resistive state or off state is almost parralel to the x-axis where as the on state has a steep slope.

Conclusions

The results support the claim that this material can be used as a memory device once a few obstacles are overcome. From the characterization of the thin film a couple conclusion can be drawn.

- After vacuum thermal evaporation, elemental sulfur occurs with in the thin film.
- Photocondensation does not affect the occurrence of elemental sulfur within the thin film of AgGeS after vacuum thermal evaporation
- Annealing does affect the occurrence of elemental sulfur within the thin film of AgGeS after vacuum thermal evaporation, and causes the sulfur to diffuse back into the bulk.
- From the image obtained from the scanning electron microscope, two phases are present, one electrically conductive the other insulative.
- The conductive material phase is on the magnitude of 50nm after deposition.

Members Roles

Carter – 4th year EE that is here to help support and give expertise to the project. Helps with depositions and trouble shooting the evaporator. Made a program to automate device testing.

M.J. – Graduate EE student that is using this projeet as his masters thesis. Does depositions, trouble shoots the evaporator, and will design the circuit for the memory device and eharaeterize the deviees.

Porter – Graduate EE student that is using this projeet as his masters thesis. Does depositions, trouble shoots the evaporator, and will design the circuit for the memory device and eharaeterize the devices.

Nieholas – 3rd Helps with deposition of deviees and characterizing them eleetrically. Helps eoordinate time on AFM and SEM and helps interpret results. Also designed a photomask inorder to fabrieate shadow masks for deposition.

References

[1] M.N. Kozieki, Mira Park, and Maria Mitkova, "Nanoscale Memory Elements Based on Solid-State Eleetrolytes," in IEEE Transaetions on Nanotechnology, vol. 4, pp. 331-338, 2005.

[2] M. N. Kozieki, M. Mitkova, J Zhu, and M. Park, "Nanoseale Phase Separation in Ag-Ge-Se Glasses," Mieroeleetronie Engineering, vol 63, pp. 155-159, 2002.

[3] M. N. Kozieki, M. Mitkova, M. Park, M. Balakrishnan, and C. Gopalan, "Information Storage Using Nanoseale Eleetrodeposition of Metal in Solid Eleetrolytes," Superlattices and Mierostruetures, vol. 34, pp. 459-465, 2003.

[4] M. N. Kozieki, M. Yun, S. –J. Yang, J. P. Aberouette, and J. P. Bird, "Nanoscale Effeets in Devices Based on Chaleogenide Solid Solutions," Superlattices and Mierostruetures, vol. 27, no. 5/6, 2000.

HONORS PROJECT

Solar Powered Wireless Access Point

Projeet Investigators:

Julian Corwin and Derek Lockhart
(John Seng-Advisor)
Computer Science
California Polyteehnie State University
San Luis Obispo, CA

Executive Summary

The solar powered wireless aceess point projeet explores the possibilities of using renewable energy resources to provide off-grid wireless access to remote areas. High expenses in the implementation of wireless networks occur when the desired coverage area lacks the infrastructure to support the wireless access point. Standard wireless access points require a power conneetion and an Ethernet connection.

The Solar Powered Wireless Aeeess Point (WAP) research projeet seeks to eliminate these infrastrueture needs in developing a remote power generation and wireless signal extension system. Development of this system would allow a single internet connection servicing few users to extend into a broad eoverage area servicing many.

Why solar powered?

Solar power offers unique advantages in comparison with other renewable energy sources. Solar panels are lightweight, eontain no moving parts, are highly mobile, and generate reliably and predictably. Using solar panels allows the construction of a single mobile unit, eapable of generating and storing power while providing wireless internet access.

Who would use this?

Countries without extensive electrical and internet infrastructure would find this product useful in bringing the internet to rural communities. Cities with free WiFi programs, and organizations such as the One Laptop per Child Association would find a solar powered access point an eeonomieal and practical approach for internet eonnection.

Introduction

This report outlines the researeh completed by Derek Lockhart and Julian Corwin during the winter and spring of 2007. Background information on the basics of solar panel power generation, battery energy storage, and the electronics uscd in this project are included for the purpose of edueating future students who eontinue this projeet.

The end goal of the project is to build a functional system capable of accepting multiple input power sources. The project team intends to build a simple charging system comprised of the Micro-controller, charge controller, battery, solar panel, wireless access point, and switching regulator, and develop additional power sources as time allows. These components will complete the objective of building a "proof of concept" system that supplies power to the wireless access point, and recharges using renewable energy sources. Beyond the proof of concept system additional renewable power sources will be made compatible with the system. The system should contain the following features:

- Produce energy from renewable sources
- Charge the stored energy
- Operate a wireless aeeess point to extend network coverage
- Implement a microcontroller to maximize effieiency

These features were approaehed from a system level initially with purehasing equipment to fulfill speeifie needs of the projeet, and to speeifie eleetronie components when neeessary. Equipment was purehased to minimize eost and maximize funetionality, as the wireless aeeess point should have a priee range competitive with eurrent wireless implementations.

Background

This seetion elaborates on the details of the eomponents used in this project, including basie energy and power calculations, and brief definitions of the equipment used.

Energy and Power Calculations

Understanding the basies of power generation and storage requires some working knowledge of Eleetrieal Engineering.

Power Consumption

The average power consumed or produeed by electrieal devices is deseribed by the following equation:

$$P = IV$$

Where:

- P is the power generated in eonsumed in watts (joule/second).
- 1 is the current in amperes
- V is the voltage in volts

Energy Storage

The energy stored by a battery ean be determined from the battery's amp-hour (AH) and voltage ratings. All batteries are rated on standards of voltage and amp-hours, and thus the energy stored in them ean be determined. A battery will be rated with x amp-hours, meaning the battery is eapable of supplying x amps for 1 hour or x/y amps for y hours. For example, a 12V battery with a rating of 17AH ean supply 17 amps for 1 hour, or 1 amp for 17 hours. The energy stored in 12V 17AH ean be found using the equation for power and the total seconds in an hour:

$$E = P * T = (12V)(17A)(1hour) = (12V)(17A)(3600\textbf{sec} / \textbf{hour}) = 734{,}400\,joules$$

Switching Regulator – A voltage regulator that uses a switehing element to transform the supply into an alternating eurrent, whieh is then converted to a different voltage using capaeitors, inductors, and other elements, then converted baek to DC. The circuit ineludes regulation and filtering components to insure a steady output. Advantages inelude the ability to generate voltages beyond the input supply range and effieieney; disadvantages inelude complexity and eost. (1)

Important aspeets of a switehing regulator are:

- Input voltage range
- Output voltage and eurrent
- Effieieney

Solar Panel – A photovoltaic deviee that eonverts sunlight into electricity. Solar panels are deseribed by three parameters: open eireuit voltage, short eireuit eurrent, and maximum output power. Open circuit voltage is the voltage aeross the positive and negative terminals of the solar panel in full sunlight. Short eireuit current is the current produeed when the positive and negative terminals are eonneeted. Maximum output power of solar panels is measured at the "knee" of the power curve, at an output current and voltage that maximizes power, as illustrated in figure 1.

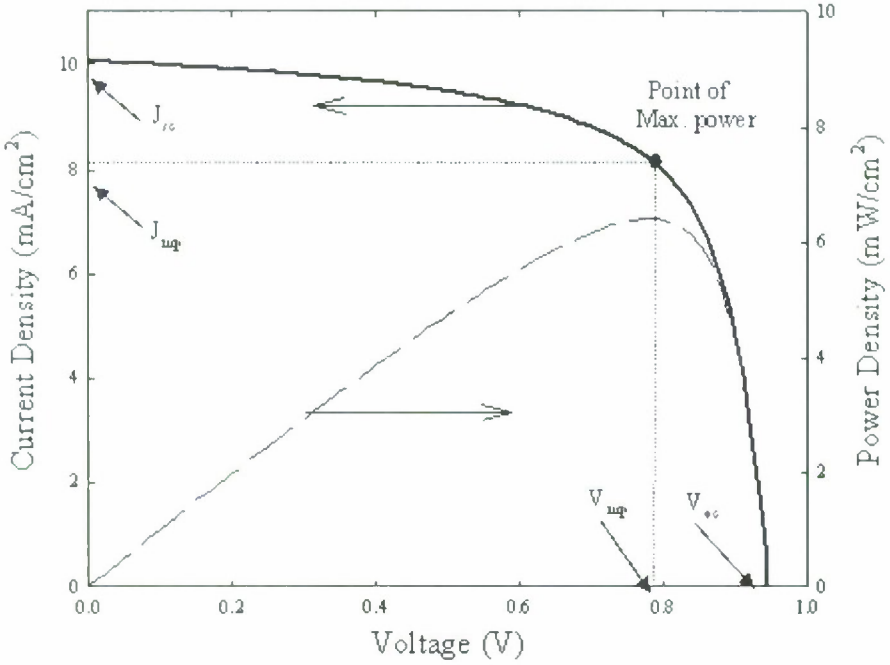


Figure 1 - Power characteristics of a solar panel (2)

Charge Controller – Controls the power flow between a solar panel and a battery. The charge eontroller ensures the battery does not over eharge and provides isolation of the solar panel from the battery.

Design

Goal: To create a system that aeeepts multiple sourees of input power, stores the input power, and converts it to a form usable by an off-the-shelf wireless range extender.

The following figure represents a generie system level design eapable of accepting multiple input power sourees, storing that power, and delivering that power in a form aeeepted by the wireless router or range extender.

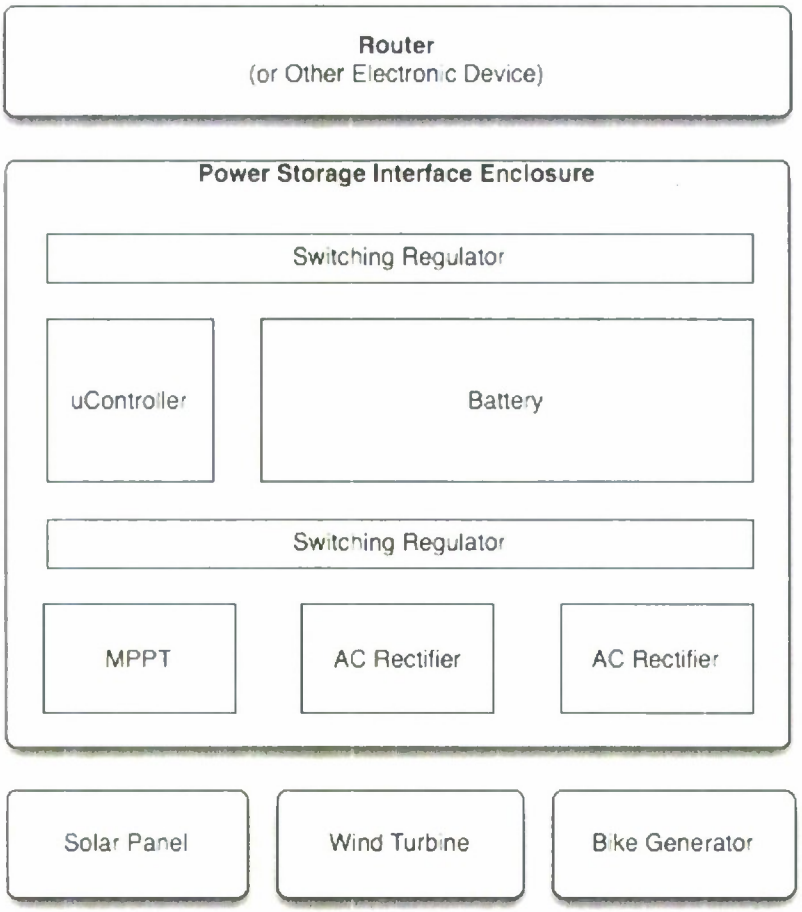


Figure 2 – Wireless Access Point with multiple power sources

Components

The following components were purchased or acquired for the Solar Powered Wireless Access Point projeet:

Solar Panel

A 5W, weatherproof solar panel was available from a previous research projeet and utilized in this design. Seleetion of a larger panel would allow the wireless aeeess point to operate constantly, however, for proof of design concept this solar panel proved to be adequate.

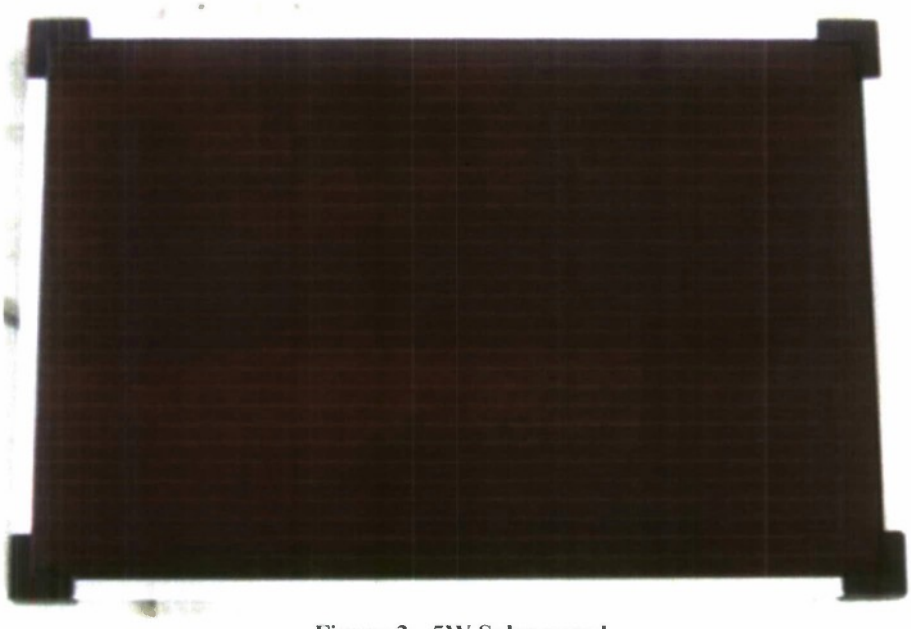


Figure 3 - 5W Solar panel

Deep Cycle Battery

A 12V 17AH deep eycle battery was seleeted based on its cost and ability for deep discharge and reeharge eyeles. This battery operates at normal car battery voltage, however ear batteries do not exhibit the same diseharge eharaeteristies of a deep eycle battery and have shortened life-spans after undergoing deep-cycle discharges.

Figure 4 - 12V, 17Amp-hour deep cycle battery

Charge Controller

The SunGuard 4.5 amp 12 volt solar charge controller was selected for use with this project. Most charge controllers are in use for relatively large solar arrays, which explains the high power capacity of this charge controller (P = IV = (4.5)(12) = 54watts). This charge controller was selected for its minimal cost, size, and compatibility with our solar panel and battery.

Figure 5 - Sunguard charge controller

Switching Regulator

With the battery operating and eharging between voltages of 12 and 14volts, a voltage regulator eapable of handling a wide range of inputs was required. After searching numerous electronic component sites, the C&D technologies 10W non-isolated, single output DC-DC eonverter was selected for the following reasons:

- Wide input range (-7 to 28VDC)
- Short circuit protection
- Effieiencies up to 95%

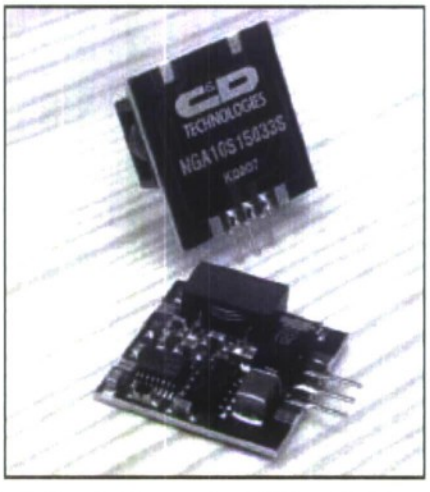


Figure 6 - C&D Technologies 3.3V switching regulator

Methods

Energy Requirements

The energy requirements of the wireless range extender had to be determined in order to estimate the required size of the solar panel and battery system. Initial estimates for power consumption of the wireless range extender were between 5-10 watts. After purchasing of the Belkin Wireless G Universal Range extender (figure 2) a Kill-a-Watt meter (figure 3) was connected and the power consumption was measured.

Figure 7– Belkin range extender

Figure 8 - Kill-a-watt power meter

Power consumption was measured to be 5 watts after connecting the transformer to the Kill-a-watt meter. However, the transformer introduces inefficiencies in the conversion of AC wall power to the DC input power required by the range extender. The casing to the Belkin range extender was removed, revealing the power line input circuitry. This circuitry consisted of RF chokes, decoupling capacitors, and a linear regulator. The linear regulator reduced the 5V output of the transformer to the 3.3V required by the internal circuitry, introducing a power loss:

$$P_{loss} = (inputvoltage - outputvoltage)(inputcurrent) = (5 - 3.3)(180mA) = 300mW$$

The wireless access point was then unplugged from the transformer, and the power flowing just through the transformer was measured at 1 watt, indicating that the transformer itself introduced power loss from energizing the transformer coils. The energy conversion process for a router connected to a standard AC wall plug is shown below:

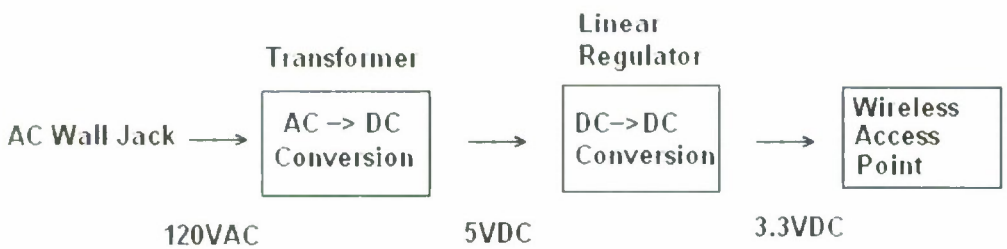


Figure 9 - Energy conversion process for wall-plugged router

The transformer and linear regulator blocks introduce power losses in the conversion process; however, since the router must receive power from a 12V DC battery the transformer will not be necessary. However, the linear regulator inside the wireless

aecess point cannot aceept input voltages above 8V, even if it were the power loss from the linear regulator would be even greater than before:

$$P_{loss} = (inputvoltage - outputvoltage)(inputcurrent) = (12 - 3.3)(180mA) = 1.56watts$$

To solve this problem the internal regulator was removed and replaeed with the C&D technologies switehing regulator, yielding a much more direct energy conversion.

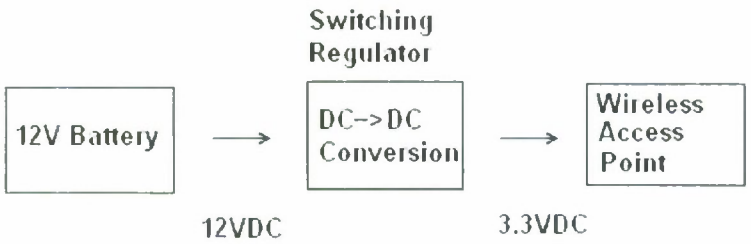


Figure 10 - Conversion process with switching regulator

To conneet the aceess point to the battery voltage, the transformer eable was severed and soldered to two alligator elips, allowing eonnection of the aceess point to the 12V battery. The battery output voltage and eurrent were measured at 12.7V and 180mA respectively, yielding a power consumption of:

$$P = IV = (12.7)(180mA) = 2.3watts$$

A 54% decrease from the power consumption measured when the aceess point is connected to standard AC power. The aceess point with the replaced regulator is shown in figure 10.

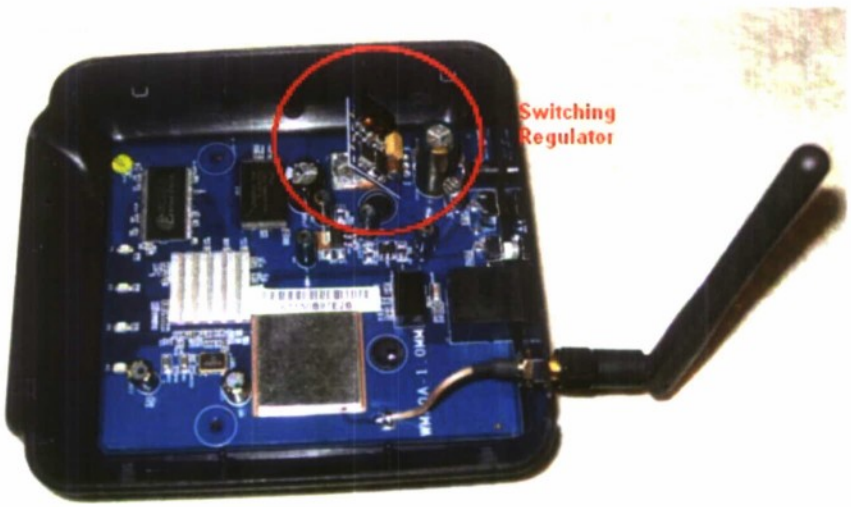


Figure 11 - Access point with switching regulator in place

Solar Powered Wireless Access Point

With the aecess point able to run from off battery power, the system components were assembled into a weatherproof box, shown in figure 12.

Figure 12 - Access Point, charge controller, and battery in weatherproof casing

The solar panel was eonneeted to the eharge eontroller to eharge the battery, completing the solar powered wireless aeeess point system, as shown in figure 13.

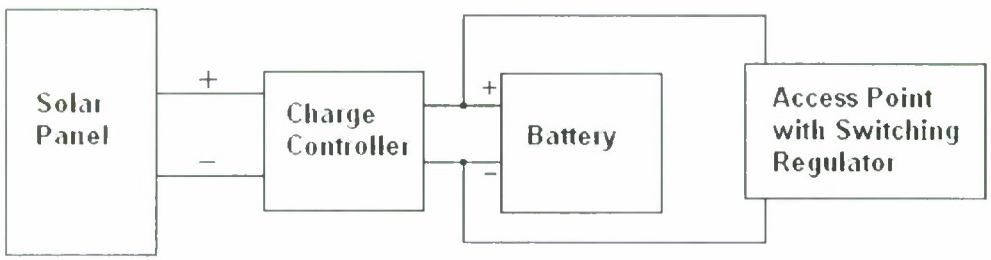


Figure 13 - Solar Powered Wireless Access Point System Diagram

Bike Generator

During spring quarter the projeet team deeided to pursue alternatives to solar power. A bike trainer was purehased by Dr. Seng at the eost of $8 at a garage sale, and fitted with a used alternator from a 1992 Honda Civie station wagon. Inelusion of the bike generator allows for more variability in the size of the solar panel. The bike generator produeed 20 watts of power after surpassing a threshold RPM rate. The sehematie for the alternator, as shown in a Honda ear manual is shown in Appendix A, however this schematie differs slightly from the alternator used for the bike generator.

Alternator Pins

Bolt Terminal – This is the output voltage terminal. Output varies from 14-16V and 0.3-2.5A depending on the state of the battery. When the battery is nearly empty, the alternator will produce more power, scaling down as the battery reaches full charge.

Black/Yellow – Requires battery voltage to operate the field circuit. Without battery voltage on this terminal the alternator will produce no output current or voltage.

Yellow/White – This pin maintains a voltage of 8V when the alternator is in operation, and sits at a voltage slightly lower than 8V when the alternator is not in use but the field circuit is charging. Grounding this terminal reduces the alternator output to zero.

Simulations

Simulations of the power system were created using Matlab. The simulations take into account the following variables:

- Input power of the solar panel
- Power requirements of the access point
- Battery capacity
- Cycling time
- Duty cycle of the access point
- Hours of sunlight

These variables are used to plot the energy stored in the battery versus time in days. If the user rides the bicycle, this added input power is displayed in the first portion of the day as an increased slope of duration proportional to cycling time. The simulations were run twice, once without the bicycle, and once with a 20 second cycling time at 20 watts. Appendix B contains the matlab simulation code

The simulations were run with the following assumptions:

- 4 watts of input power
- 2.5 watts of power required for the access point
- 10 hours of sunlight
- 20 watts of power generation from bicycle

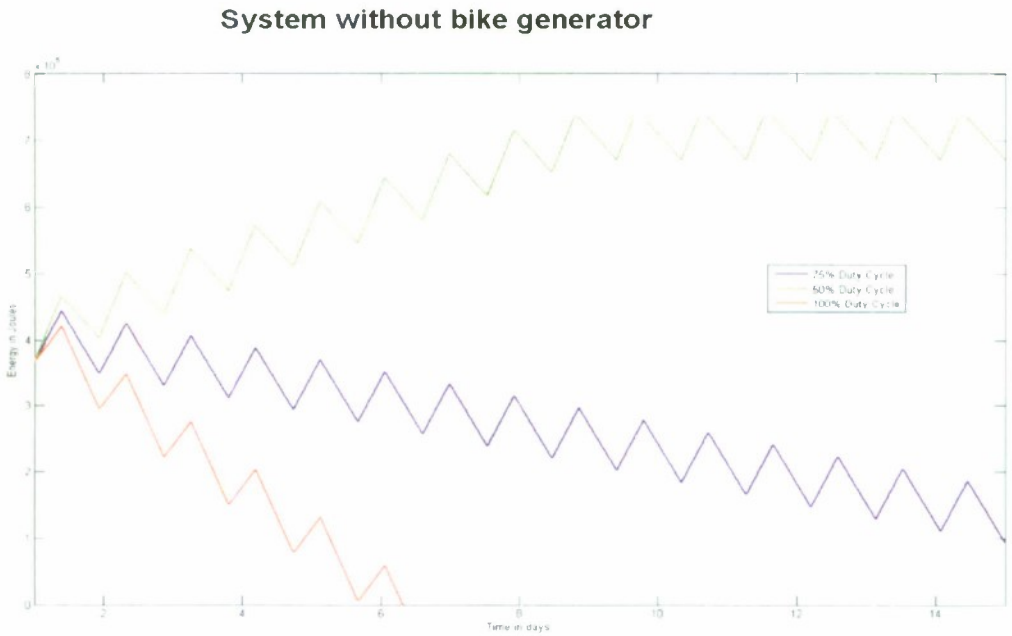


Figure 14-Variable duty cycle, solar panel input only

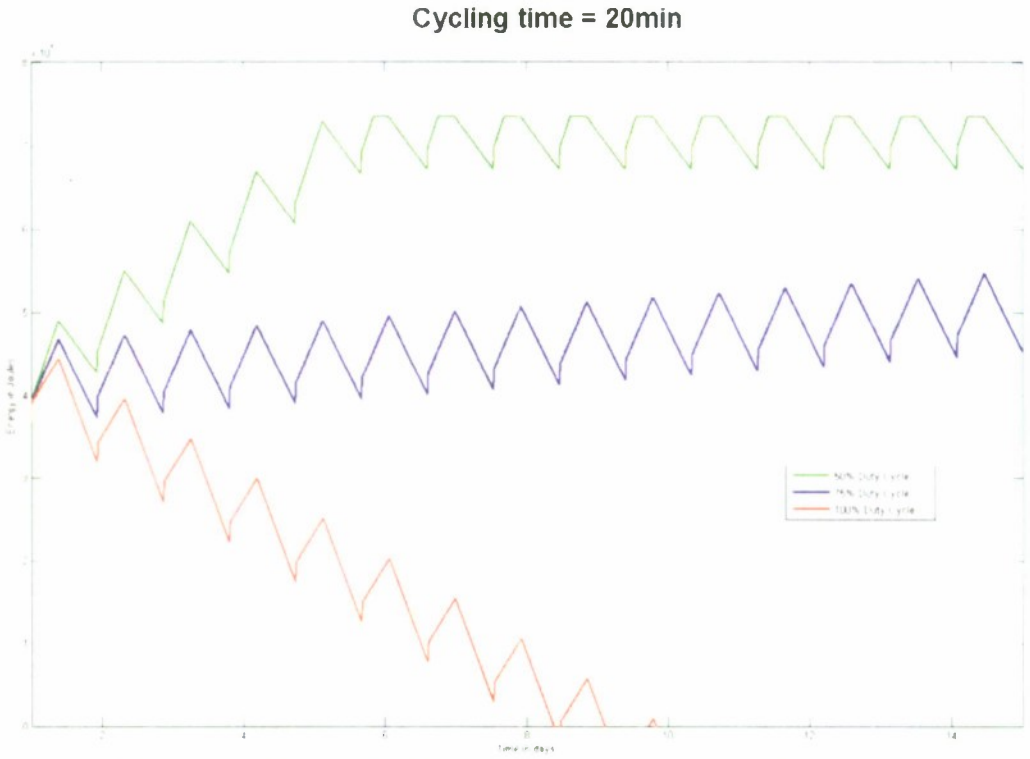


Figure 15 - Variable duty cycle, solar panel input and 20 minutes of bicycle input

Conclusions

The solar powered wireless aceess point set out to develop a feasible solution to wireless internet aeeess using renewable energy resourees. Initial goals were:

- Produee energy from renewable sources
- Charge the stored energy
- Operate a wireless aceess point to extend network coverage
- Implement a mieroeontroller to maximize effieiency

The projeet team eompleted all these goals exeept the mieroeontroller implementation. The charge controller system, alternator, and solar panel all functioned without the need for intelligent system eontrol, and so the mieroeontroller is unneeessary. However, future projeets should eonsider implementation of a mieroeontroller as a means of providing a user interfaee of input and output power. Intelligent eontrol of the aeeess point eould provide duty eyele sealing proportional to the input power, thereby preventing system failure.

We reeommend future research teams foeus their efforts on the following aspects of the projeet:

- Implement a microcontroller user interfaee to eontrol and display system parameters sueh as:
    - battery eharge state
    - power input
    - power output
    - daily power generation
    - bieyele power generation
- Implement a new generator to replaee the alternator. The replacement should have the following features:
    - No battery voltage requirements
    - Compatibility with eurrent battery
    - Output eurrent proportional to eyeling speed

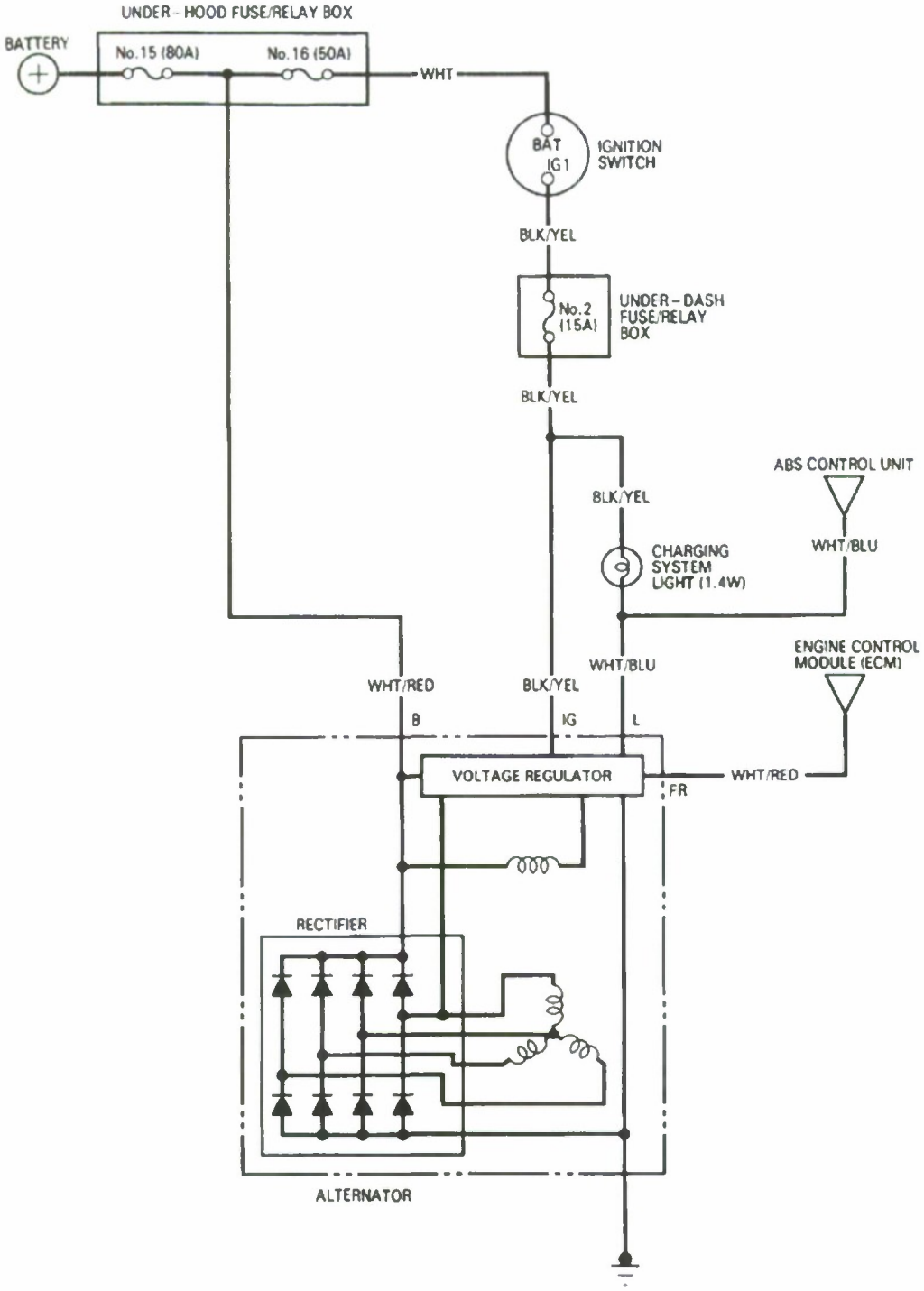
These features will make the solar powered wireless aecess point a more user friendly device, and allow for easier implementation. Furthermore, the addition of a new bieycle generator to replaee the alternator would allow more effieient bicyele input appealing to areas with variable weather eonditions.

Bibliography

1. "Switching Regulators" *MAXIM, Analog, Linear, and mixed signal devices.* Accessed on 6/10/2007 <http://www.maxim-ic.com>

2. Pearce, Joshua M. PhD. *Solar Cell Research.* Accessed on 6/9/2007 <http://jupiter.clarion.edu/~jpearce/pv_files/image012.gif>

Appendix A – Alternator Pin Description



Appendix B – Matlab Simulation Code

```
function [] = psim()
Pin=input('Enter the input power of the system: ');
Po=input('Enter the power consumption of the router: ');
sunlight=input('Enter the effective hours of sunlight: ');
Dutycycle=input('Enter the router Duty Cycle: ');
biketime=input('Enter the amount of cycling time per day in minutes:
');

days = 15;
Emax = 12*17*60*60;
totalminutes = days*24*60;
minutes=0;
energy=Emax/2;
hours=0;
for i = 1:totalminutes
    minutes=minutes+1;
    hours=hours + 1/60;
    if (hours <= sunlight)
        if ((hours >= 0) && (hours <= (biketime/60)))
            energy = energy + 60*(Pin - Po*Dutycycle)+60*20;
        else
            energy = energy + 60*(Pin - Po*Dutycycle);
        end
    else
        energy = energy - 60*Po*Dutycycle;
    end

    if (energy > Emax)
        energy = Emax;
    end

    if (hours >= 24)
        hours=0;
    end


    eotime(i)=energy;

end

plot(linspace(1,days,(totalminutes)),eotime)
xlabel('Time in days')
ylabel('Energy in Joules')
xlim([1 days]);
ylim([0 8e5]);
```

Appendix C – Cost

Solar Components	**Price**
Belkin Wireless Range extender	$39.99
Deep Cycle 12V 17AH Battery	$28.60
Charge Controller	$28.50
5W Solar Panel	$49.99
Switching Regulator	$24.44
Weatherproof casing	$39.99
Total	$211.51
Bicycle Generator Components	
Alternator	$35.00
Belt	$10.00
Bike trainer modifications	$25.00
Bike trainer modifications	$9.00
Total	$79.00

Appendix A

Research Forum Announcement

CAL POLY

C3RP
CALIFORNIA CENTRAL COAST
RESEARCH PARTNERSHIP

High-Tech Industry Breakfast Forum

Date:	February 20, 2009
Time:	7:30 am – 9:00 am
Location:	Poly Canyon Village, Aliso Conference Room A007-A014 Cal Poly Campus

You are cordially invited to participate in our on-going breakfast forum for the local high-tech industry, hosted by Dr. Robert Koob, Provost and Vice President for Academic Affairs, and Dr. Susan Opava, Dean of Research and Graduate Programs, at Cal Poly. The purpose of this forum is to bring together CEOs and Chief Technology/Engineering Officers from high-tech companies in the region, on an on-going basis, to foster a mutually beneficial relationship between the university and this important local industry sector.

7:30 am	Continental Breakfast - Small group discussions & networking
7:50 am	Introductions and Overview of Program
8:05 am	Poly Canyon Village Project - Alan Pepe, Associate Director Housing/Director Housing & Business Services
8:15 am	Cal Poly Highlight: Algal Production and Biofuel Development Programs
8:40 am	Company Highlight: Phycotech Corporation – Paul Cathcart, CEO
9:00 am	Adjourn

Please RSVP # of participants to: Dejana Lubura 805-756-5153 or dlubura@calpoly.edu by **February 11, 2009**

Appendix B

Relevant Publications

Available online at www.sciencedirect.com

ScienceDirect

Optics and Lasers in Engineering 45 (2007) 1067–1070

OPTICS and LASERS ENGINEERING

www.elsevier.com/locate/optlaseng

A phase-stepped grating technique for frequency shifting in laser Doppler velocimetry

J.P. Sharpe

Department of Physics, Cal Poly State University, San Luis Obispo, CA 93407, USA

Available online 12 July 2007

Abstract

We demonstrate a novel approach to the moving grating technique for removal of velocity ambiguity in a laser Doppler velocimeter. The method replaces the continuous motion of a moving grating with a grating which can be positioned or switched to three distinct phase positions. Advantages of this approach are its potential for compact implementation and high stability of the velocity offset.

Keywords: Laser Doppler velocimetry; Phase-stepping

The technique of laser Doppler velocimetry (LDV) has been in development for some forty years and still represents one of the best methods for high-accuracy, single-point, fluid velocity measurement [1]. A recurrent challenge is resolution of the ambiguity in the direction of the fluid flow. One of the most common techniques to resolve this ambiguity is a Bragg cell which is used to impose a frequency shift on one or both beams of the LDV system. However, Bragg cells have disadvantages since they can be misaligned through mechanical shock and they require a bulky power supply. Because of their size they are also not suitable for miniaturization for inclusion in compact LDV probes and while delivering the light through optical fibers is a possible solution this also has problems arising from differential strain in the optical fibers. A recent review and discussion of frequency shifting techniques can be found in [2].

One of the earliest methods of making the LDV system direction-sensitive was by the incorporation of a moving diffraction grating in the optical path and this is schematized in Fig. 1. The diffraction grating fulfills two roles in this system. First, it splits the input laser beam in two beams (if a sinusoidal phase grating is used only two beams are generated otherwise all but the ± 1 orders are blocked). Secondly, through its motion, it induces a Doppler shift onto the two diffraction orders, shifting one order up in frequency and the other down. When the beams recombine at the measurement volume and scatter from a particle in the flow, the light from both beams beat against each other at the Doppler-induced difference in frequency. This beat frequency is superimposed on the Doppler shift due to the particle motion and, depending on whether the motion causes an increase or decrease of the beat frequency, allows the flow direction to be determined. In practice the grating is reproduced radially on a small glass wheel mounted on a motor and rotates continuously [3,4]. Although this method of frequency shifting works quite well, and is employed by different manufacturers of LDV systems, it has several drawbacks. These include a spatially varying fringe period in the measurement volume, vibration which is induced by the motor itself and the need for a feedback system to ensure that the grating is rotating at a constant rate.

Here we present a variation of this moving grating technique where we note that, firstly, it is not necessary for the phase position of the grating to increase continuously but rather need only vary from 0 to 2π radians and, secondly, that only three discrete steps are needed in the interval $[0\ 2\pi]$. The use of discrete phase steps has been known for some time in the field of phase-shifting interferometry [5] and the rationale for this method in the context of LDV can be appreciated by noting that the use

E-mail address: jsharpe@calpoly.edu

0143-8166/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.optlaseng.2007.05.005

of a moving diffraction grating to impose frequency shifts on the ± 1 diffraction orders in order to give moving fringes is the same as projecting a (spatially filtered) image of the moving grating into the flow. In order to accomplish one complete cycle of the image the grating just has to be moved so that one period of the projected fringe pattern is covered. Once this is done the grating image is reset to its initial position and the motion repeated. A schematic of phase-stepping is shown in Fig. 2 where the grating is sequentially moved by one third of a period. After each movement the grating "dwells" for a time $T/3$ where T is the period of the motion and at the end of one period is reset back to its initial position.

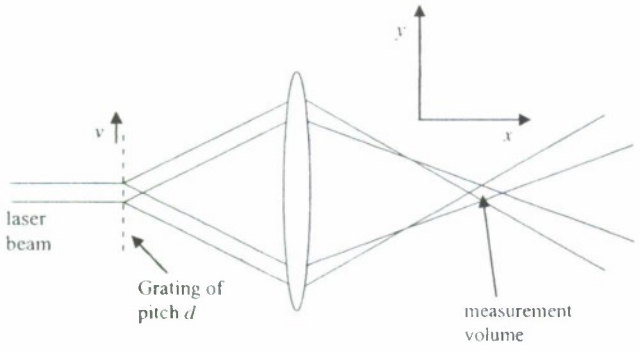


Fig. 1. Schematic of an LDV system with a moving grating. Light from a laser is diffracted into two beams using the grating (for clarity only the ± 1 diffraction orders are shown). The grating moves transverse to the beam with a velocity v so that the diffraction orders are shifted up and down in frequency. The diffracted beams are then incident on the lens and where they cross defines the measurement volume. The fringes in the measurement volume move in the y direction with a frequency $2v/d$.

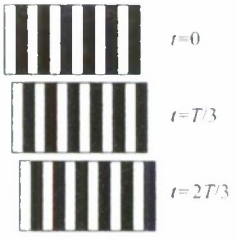


Fig. 2. Phase-stepping a grating. The three images show the grating position at the times indicated. If the desired frequency shift is $1/T$, the grating is first positioned as shown for $t = 0$. It is then moved (effectively instantaneously) at times $t = T/3$ and $t = 2T/3$. At time $t = 3T/3$ the grating position is reset to the position shown for $t = 0$. Then the whole process repeats.

In order to demonstrate the concept we use an amplitude grating mounted on a piezo-electric translation stage which can accurately step the grating to three discrete positions. Although we use a piezoelectric transducer in this work, the frequency limitations of such devices (a few kHz for the fastest that are commercially available) would preclude their use for all but the slowest flows. Later we will discuss some of the methods that could be used to implement the technique at much higher frequencies.

The experimental arrangement is shown in Fig. 3 where light from a low power HeNe laser ($\lambda = 633$ nm) is incident on a 40 lp/mm Ronchi grating (25 μm pitch). The ± 1 orders are selected and the beams crossed on the ground glass screen with a grating magnification of ~ 1. Forward scattered light from the screen is collected and imaged through a 50 μm diameter pinhole and detected with a phototube. The optical intensity fluctuations are recorded and Fourier transformed using a digital oscilloscope.

The grating is mounted on a PI P733.2CL piezo stage and the stage driven with a PI E503 LVPZT amplifier. The frequency response of this system is limited to several hundred hertz and we run the system in "open loop" (which means we disable the feedback from capacitative position sensors in the piezo stage) to achieve reasonably high frequencies. Since the fringes are formed using the ± 1 orders of the grating, a shift of half the period of the grating is needed to translate the fringes by one period, which is 12.5 μm. The technique calls for translations of

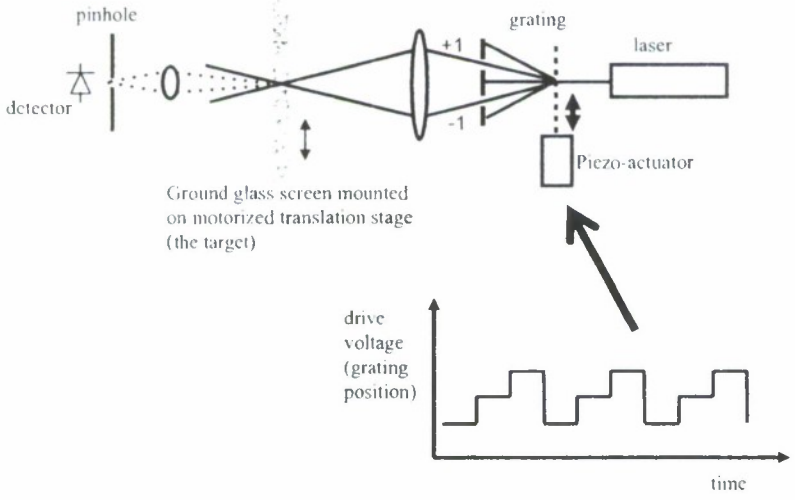


Fig. 3. Schematic of arrangement to test the concept of discrete phase stepping in a laser Doppler system.

0, 1/3 and 2/3 of the fringe period, so the positions of the diffraction grating should be 0, 4.2 and 8.3 μm. We programmed a function generator (Agilent 33120A) to supply the correct voltages to the PZT amplifier which drove the piezo stage. When driven at low frequencies the required positions could be achieved to ±0.1 μm, though at higher frequencies there were considerable amounts of ringing and overshoot. Thus, we had to drive the stage at fairly low frequencies, and we settled for 20 Hz.

The experimental data are shown in Fig. 4 and the power spectra obtained from a computer simulation are superimposed on the data. The computer simulation assumes only one particle in the measurement volume and neglects the finite extent of the fringe pattern. In this way the intensity of the scattered light as a function of time can be written [1].

$$I(t) = I_o\left[1 + \cos\left(\frac{2\pi vt}{\Lambda} + \frac{2\pi}{3}\,\mathrm{floor}\left(\frac{2\pi f_0 t}{2\pi/3}\right)\right)\right],$$

where I_o is a constant, v is the component of the particle velocity across the fringe pattern, Λ is the fringe period, and f_0 is the frequency at which we drive the grating. The operator floor rounds its argument to the nearest integer in the direction of negative infinity. Note that the second term in the argument of the cosine is the stepped version of a continuous phase shift.

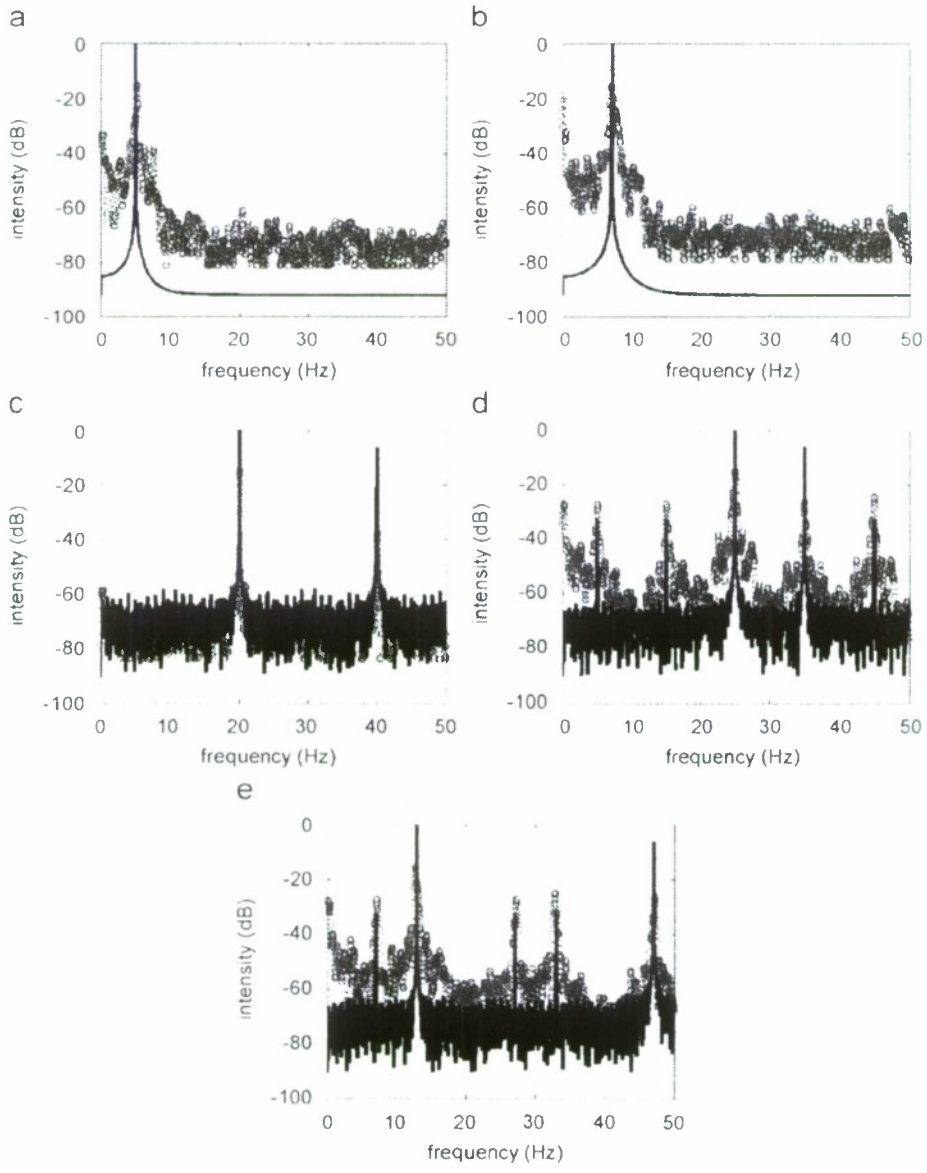


Fig. 4. Data from the experimental arrangement (o) and the computer simulation (–) (a) Ground glass screen driven forward. The Doppler peak is at 5 Hz. (b) Ground glass screen reversed. The Doppler peak is at 7 Hz. (c) The ground glass screen is held stationary while the grating is oscillated at 20 Hz, leading to a peak at 20 Hz and a harmonic at 40 Hz. (d) Ground glass screen driven forward and grating oscillated, predominant peak at 25 Hz. (e) Screen direction reversed and grating oscillated, predominant peak at 13 Hz.

Fig. 4(a) shows the spectrum when the target (the ground glass screen) is driven forward, yielding a spectral peak at 5 Hz. In this case there is no movement of the grating. In Fig. 4(b) the screen is driven backwards and the peak frequency occurs at 7 Hz. The difference in peak position due to forward and reverse motion is because our available translation stage moves slightly faster in the reverse direction than in the forward direction, but there is no way to tell a priori the direction of the travel. In Fig. 4(c) the ground glass screen is stationary and the grating is oscillated at 20 Hz so there is a distinct peak at 20 Hz. Because of the quantized nature of the grating position we can also see a strong harmonic at 40 Hz. In Fig. 4(d) the ground glass screen is driven forward and the grating is oscillated. The predominant peak is at the sum of the frequency due to the motion (5 Hz) and the frequency due to the oscillating grating (20 Hz) which is 25 Hz. This peak stands about 10 dB above the next highest peak in the spectrum. One can also see that the results of the computer simulation not only show peaks at the same position as in the experiment, but that the peaks have approximately the same relative strengths. Fig. 4(e) shows the case when the screen is reversed and the grating oscillated. Now the largest peak occurs at 13 Hz which is the difference between the frequency due to the oscillating grating (20 Hz) and the frequency due to the motion (7 Hz).

Because of the strong harmonics introduced by the discrete nature of the phase stepping, there are many spectral peaks in the signal (Fig. 4(d) and (e)). As long as these are well separated from the dominant signal peak there should be no problem in making valid velocity measurements. However, the extra peaks limit the range of velocities that can be measured. In particular, the velocity of the flow must not lead to frequency shifts that are greater than or equal to half the frequency which is induced by the moving grating.

In summary, we have demonstrated what we believe to be a novel approach to the "rotating grating" technique for removing the velocity ambiguity in a laser Doppler velocimeter. The method requires a grating that can be stepped discretely in phase. There are several high-speed technologies that suggest themselves for implementation of this technique. These include liquid crystals, which can have sub-microsecond switching speeds [6] and micro-electrical mechanical systems (MEMS), such as those used in the Texas Instruments DLP devices, which have switching times in the tens of microseconds [7]. Thus the method could achieve frequency offsets of order 100 kHz using currently available technology. One benefit of implementing this technique with an electrically switched grating is that the frequency shift is only determined by the timing of the drive electronics and this can be done with high precision. In the case of compact laser Doppler probes this technique would also be advantageous since an electric motor would not have to be accommodated within the probe head. In contrast to Bragg cells this technique is fairly robust since minor misalignments of the optical system will not affect the diffracted beams; however with current technology our approach can not provide the high frequency shifts (~10 MHz) required for very high speed flows.

Acknowledgements

This work is supported by the Department of the Navy, Office of Naval Research, under Award no. N00014-06-1-1111.

References

[1] Durst F, Melling A, & Whitelaw JH. Principles and practice of laser Doppler anemometry. London: Academic Press; 1976.
[2] Czarske JW. Laser Doppler velocimetry using powerful solid-state light sources. Meas Sci Technol 2006;17:R71–91.
[3] Stevenson WH. Optical frequency shifting by means of a rotating diffraction grating. Applied Optics 1970;9:649–52.
[4] Oldengarm J, Venkatesh P. A simple two-component laser Doppler anemometer using a rotating radial diffraction grating. J Phys E 1976;9:1009–12.
[5] Creath K. Temporal phase measurement methods. In: Robinson DW, Reid GT, editors. Interferogram Analysis. Bristol: I.O.P Publishing; 1993.
[6] Clarke NA, Lagerwall ST. Submicrosecond bistable electrooptic switching in liquid crystals. Appl Phys Lett 1980;36:899–904.
[7] Duncan W, et al. DLP switched blaze grating: the heart of optical signal processing. Proc. SPIE, 2003;4983.

J. Coat. Technol. Res., 4 (1) 107–110, 2007
DOI 10.1007/s11998-007-9002-7

BRIEF COMMUNICATION

Shear-induced 1-D alignment of alumina nanoparticles in coatings

Lucas J. Brickweg, Bryce R. Floryancic, Erik D. Sapper, Raymond H. Fernando

Abstract Atomic Force Microscopy and Scanning Electron Microscopy were used to study shear-induced alignment of alumina and silica nanoparticles in two-component polyurethane clear coatings. 1-D strings of nanoparticles, formed in an extended pearl-necklace fashion were observed near the surfaces of cured films at nanoparticle volume fractions less than 0.05. This alignment is affected by the shear conditions of the application method. When applied by spraying, linear particle strings as long as 5 cm were observed in the direction of shear. Nanoparticle strings were also found, to a lesser extent, when coatings were applied by a drawdown method. The phenomenon was not observed in coatings applied with minimal shear. These particle string formations, in addition to affecting the performance of coatings, may have broader implications in the field of nanomaterials. Our literature searches so far have not uncovered reports of stable, 1-D nanoparticle arrangements with same degree of linearity produced under simple shear with compositions having very low particle loadings.

Keywords Nanoparticles, Nanocomposite coatings, Nanoparticle ordering, 1-D nanoparticle alignment, Polyurethane coatings, Nanoparticle processing

Presented at the 2006 FutureCoat! conference, sponsored by the Federation of Societies for Coatings Technology, in New Orleans, LA, on November 1–3, 2006.

L. J. Brickweg, B. R. Floryancic, E. D. Sapper, R. H. Fernando (✉)
Polymers and Coatings Program, Department of Chemistry and Biochemistry, California Polytechnic State University, San Luis Obispo, CA 93407, USA
e-mail: rhfernan@calpoly.edu
URL: www.polymerscoatings.calpoly.edu

Introduction

Experimental and theoretical aspects of highly ordered colloidal particle assemblies have received a great deal of attention in recent years.[1–9] Applications such as electronics, optics, and medical science could benefit from the unique properties of such highly ordered materials. Many of the previous studies have addressed shear-induced alignment of concentrated hard sphere suspensions. Particle size uniformity is an important factor in forming ordered structures under shear flows. Shear-induced particle string structure formation has also been reported in such dispersions, and such behavior has been correlated with shear-thinning at high rates of shear flow.[1,5] The transient string formation has been shown to occur as 3-D ordered structures of particles begin to break up leaving strings of particles aligned parallel to direction of the shear flow.[1]

The vast majority of prior literature in this area deals with 2-D and 3-D ordered structures. Literature on 1-D particle arrangements is limited, but an excellent review has been recently published.[10] The most widely reported method for producing pearl-chain formations are dielectrophoresis and other electrostatic techniques.[4,7,9] The structures created using these methods are generally complicated, and often are not single, straight wires. One common goal of arranging nanoparticles in 1-D is the production of nano-scale conductive wires for electronics applications.

This paper discusses shear-induced assembly of nanoparticles into nearly perfect straight and parallel lines produced by simple application methods. The levels of nanoparticles used are much lower than the high volume fraction studies reported in previous literature. The lines produced are single alumina and silica nanoparticles arranged in an extended pearl-necklace fashion in two-component polyurethane coatings.[11] The single particle wide linear arrangements are

Table 1: Nanoparticles Used in the Study

Type	Medium	% Solids	Particle Size
Alumina C	Methoxypropyl acetate with silicone additive	32.0	~25 nm
Alumina D	Propylene Glycol Ethyl Ether Acetate	32.0	~25 nm
Silica A	Methoxypropyl acetate/ methoxy propanol with silicone additive	32.0	~25 nm

unbranched and hold their linearity for as long as 5 cm. Our literature searches so far have not uncovered reports of 1-D straight-line arrangements of particles more than a few hundred microns in length.

Experimental

The alumina and silica nanoparticle containing polyurethane coatings were prepared according to the same methods described in a related study.[11] Commercially available alumina and silica nanoparticle dispersions (Table 1) were incorporated into the two-component polyurethane clear coating formulation at and below levels of 1.0 wt.% (i.e., approximately 0.0025 volume fraction) of the cured film.

Coating samples were prepared on glass microscope slides with different application methods and cured at 70°C for 30 min. The application methods included spray, drawdown, and gentle deposition of the liquid coating drops by a dropper. These methods represent different levels of application shear rates. Coating thicknesses were in the range of 0.25–1.5 mils (6-38 μ).

The samples were analyzed with a Pacific Nanotechnology Atomic Force Microscope using an Ultrasharp silicon cantilever's "C" tip. Close-contact mode was used for all AFM imaging. A Jasco JOEL S-8000 Scanning Electron Microscope equipped with EDS analysis capability was used to record electron micrographs and perform elemental analysis.

Results and discussion

AFM analysis and the visual appearance of all of the coatings examined in this study indicated that, in general, the nanoparticles are well dispersed. Few areas of nanoparticle aggregates can be found on all coatings tested. AFM images of most areas on the films are similar to what is reported in reference 11, Fig. 4.

Application by drawdown method

Drawdown application method produced nanoparticle strings that are very straight and continuous for a few hundred microns. AFM height images of coatings containing Alumina C (0.67 wt.%) and Alumina D (1.0 wt.%) are shown in Fig. 1. These samples were prepared at an approximate application speed of 5 cm/sec. Stable, multiple straight line formations of alumina nanoparticles were observed on both coatings. In some areas, well-defined single strings of particles were found. As the application speed increased from 1 to 10 cm/sec, the abundance and the length of particle strings increased. Similar images were obtained for silica containing samples. An SEM image of a silica containing sample is shown in Fig. 2.

Application by drop deposition method

Four drops of the coating were placed on a glass slide, allowed to flow-out and spread for 10 min, and cured under the same conditions as before (70°C for 30 min). No linear formations were found by the AFM analysis. Very few aggregates were visible with the optical microscope attached to the AFM, and they did not yield any evidence of linear particle strings.

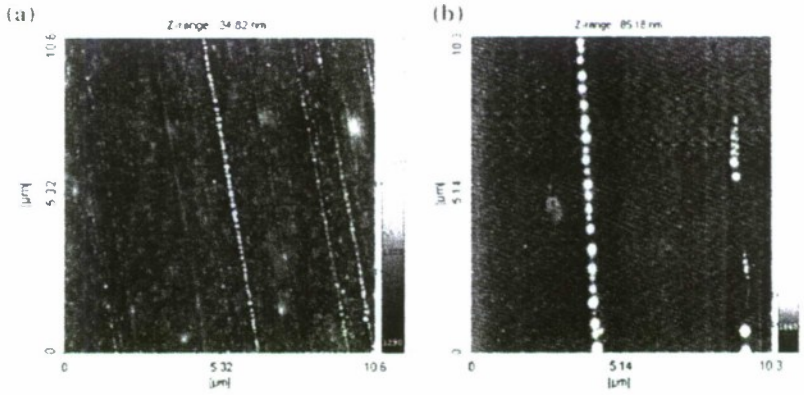


Fig. 1: AFM height images (10 × 10 micron) of coatings containing nanoparticles: (a) Alumina-C (0.67 wt.%); (b) Alumina-D (1.0 wt.%)

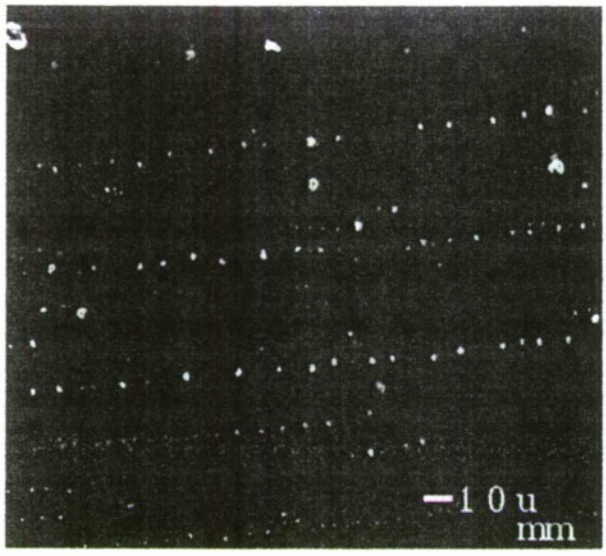


Fig. 2: Scanning Electron micrograph of Silica A (1.0 wt.%) containing coating prepared by spray application

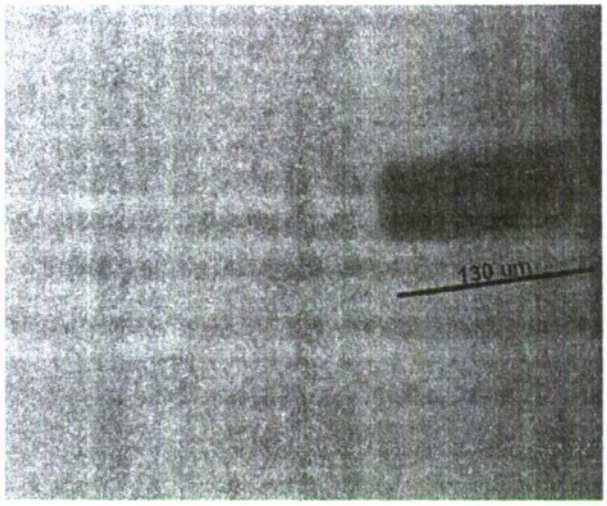


Fig. 4: Optical micrograph of Alumina C (0.5 wt.%); prepared by spray application. Linear formations exist over the entire slide parallel to the direction of application

Application by spray method

Spray application produced the largest abundance of particle strings which covered the entire microscopic slide. An SEM surface image of Silica A containing coating sample is shown in Fig. 2. AFM images of a sample containing Alumina C show many parallel strings (Fig. 3) in the spray direction. Difference in sizes appears to be due to the differences in the depth of particle strings within the film. However, the possibility exists that the larger features are nanoparticle aggregates. Detailed AFM phase contrast imaging and EDS elemental analysis by SEM are underway to determine the composition these features.

The Pacific Nanotechnology AFM used in this work utilizes a 1000× microscope with a video camera to aid in the placement of the scanning tip. Grey "streaks" can be seen on the microscope screen, and when these "streaks" or the areas surrounding them are scanned, nanoparticle strings can be found. Strings can also be seen on the microscope screen as long chains of particles, and in some cases, they cover the entire surface of the sample. Strings are visible on the optical microscope (Fig. 4), and have been measured to be more than 5 cm long with a 0.5 wt.% Alumina C sample prepared by spray method.

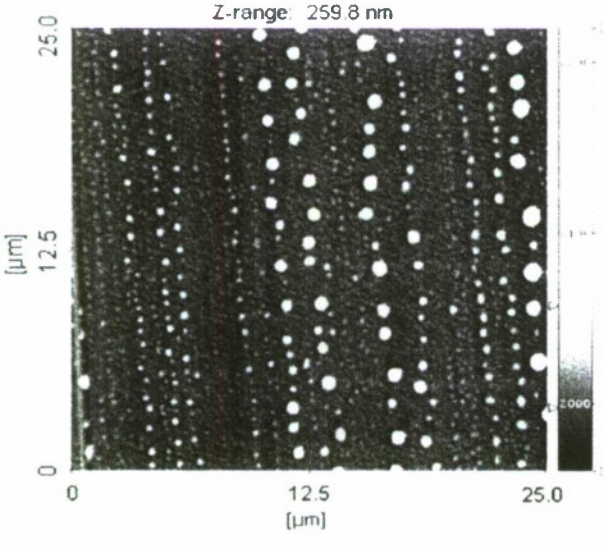


Fig. 3: AFM height image (25 × 25 micron) of coating containing nanoparticles: Alumina-C (0.5 wt.%)

Effect of shear rate

Estimated shear rates as low as 13 s^{-1} (0.05 cm/s at 3 mil wet coating thickness) produced strings of nanoparticles. Drop application method that involves minimal shear did not produce detectable nanoparticle strings. Efforts are underway to determine a critical shear rate for string formation. Spray method produced the largest abundance of particle strings. Although the modes of deformations in spray applications are complex,[7] it is a process that involves very high shear rates.

Effect of nature of nanoparticles

As indicated above, the two grades of alumina particles and one grade of silica produced the 1-D strings. It is reasonable to expect that the surface characteristics of the particles would play a critical role in this

phenomenon. The two grades of alumina are reported to have highly uniform particle sizes.[12] These grades resulted in coatings with better film clarity than others used in the study reported in reference 11. They are reported to have narrow particle size distributions. Particle size analysis of all nanoparticle samples using Transmission Electron Microscopy is underway. This analysis is expected to help to determine whether particle size uniformity is critical to formation of their 1-D arrangement.

Effects of particle strings on coating performance

The type of nanoparticle strings reported in this study could produce interesting visual effects in coatings. In addition, their distribution in this manner could produce direction dependent mechanical properties.

Future work

Further characterizations of the coatings reported in this study are underway in collaboration with the National Institute for Standards and Technology (NIST). In addition, the studies of the effects of coating compositions, viscosities, curing conditions and a wide range of other variables are also underway. Detailed particle size analysis of the alumina and silica samples is also underway.

Acknowledgments The authors would like to acknowledge Cal Poly State University's Bill Moore Endowment for financial support and BYK-Chemie Company for providing the nanoparticle samples used in this study.

References

1. Ackerson, B, "Shear Induced Order and Shear Processing of Model Hard Sphere Suspensions." *J. Rheol.*, **34** (4) 553–590 (1990)
2. Ackerson, B, Clark, NA, "Shear-Induced Melting." *Phys. Rev. Lett.*, **46** (2) 123–126 (1981)
3. Cohan, I, Mason, TG, Weitz, DA, "Shear-Induced Configurations of Confined Colloidal Suspensions." *Phys. Rev. Lett.*, **93** (4) 04600–04601 (2004)
4. Haw, MD, Poon, WCK, Pusey, PN, "Direct Observation of Oscillatory-Shear-Induced Order in Colloidal Suspensions." *Phys. Rev. E.*, **57** (6) 6859–6864 (1998)
5. Pham, HH, Gourevich, I, Oh, JK, Jonkman, JEA, Kumacheva, E, "A Multidye Nanostructured Material for Optical Data Storage and Security Data Encryption." *Adv. Mater.*, **16** (6) 516–520 (2004)
6. Kretschmer, R, Fritzsche, W, "Pearl Chain Formation of Nanoparticles in Microelectrode Gaps by Dielectrophoresis." *Langmuir.*, **20** 11797–11801 (2004)
7. Winkleman, A, Gates, B, McCarty, LS, Whitesides, GM, "Directed Self-Assembly of Spherical Particles on Patterned Electrodes by an Applied Electric Field." *Adv. Mater.*, **17** 1507–1511 (2005)
8. Shenhar, R, Norsten, TB, Rotello, VM, "Polymer-Mediated Nanoparticle Assembly: Structural Control and Applications." *Adv. Mater.*, **17** (6) 657–669 (2005)
9. Bhatt, KH, Velev, OD, "Control and Modeling of the Dielectrophoretic Assembly of On-Chip Nanoparticle Wires." *Langmuir*, **20** 467–476 (2004)
10. Tang, Z, Kotov, NA, "One-Dimensional Assemblies of Nanoparticles: Preparation, Properties, and Promise." *Adv. Mater.*, **17** (8) 951–962 (2005)
11. Floryancic, BR, Brickweg, LJ, Comer, JB, Sung, L, Forster, AM, Fernando, RH, "Effects of Alumina and Silica Nanoparticles on Automotive Clear-Coating Properties." PMSE Proceedings, ACS National Meeting, San Francisco, CA, 2006
12. Xing, L-L, Glass, JE, Fernando, RH, "Parameters Influencing the Spray Behavior of Waterborne Coatings." *J. Coat. Technol.*, **71** (890) 37–50 (1999)

2007 IEEE International Conference on Control and Automation
Guangzhou, CHINA - May 30 to June 1, 2007

Improving DC Power Supply Efficiency with Neural Network Controller

Weiming Li, Xiao-Hua Yu
Department of Electrical Engineering
California Polytechnic State University
San Luis Obispo, CA 93407, USA
E-mail: xhyu@ee.calpoly.edu

Abstract - DC-DC converters can be found in almost every power electronics device. To improve the efficiency and controller response of a DC-DC converter to dynamical system changes, neural network has been chosen as an alternative to classic methods. However, no prior work has been done in the neural network approach for control of a PSFB (Phase-Shifted Full-Bridge) converter yet. In this research, a multi-layer feed-forward neural network controller is proposed. The neural network based controller has the advantage of adaptive learning ability, and can work under the situation when the input voltage and load current fluctuate. Levenberg-Marquardt back-propagation training algorithm is used in computer simulation. The neural controller is then implemented on hardware using a DSP (digital signal processor). Satisfactory experimental results are obtained.

Keywords - neural network controller, DC-DC converter, feedforward neural networks

I. INTRODUCTION

DC-DC converters can be found in almost every electronic device nowadays, since all semiconductor components are powered by DC sources. One of the design targets for electronic engineers is to improve the efficiency of power conversion. For PWM (pulse-width modulation) converters, switching loss is an important performance measure. Many different kinds of topologies ([1], [2], [3]) have been investigated in the past to reduce the switching loss. However, those topologies either need additional components for the power circuit, which may introduce some unstable factors to the circuit, or operate at variable frequency, which makes the filter design at output stage very difficult. Phase-shifted zero-voltage switching full-bridge converters overcome the above limitations and thus have been received more and more attention recently. It employs zero-voltage-switching (ZVS) technique and allows the voltages across the transistors to swing to zero just before the start of the next conduction cycle ([4], [5], [6]).

In general, the supply voltage and load current have a wide range of variation, so the controller has to be designed to work under such conditions. The conventional approach is to assume that the converter is operated around its equilibrium state; then a set of linear equations are derived based on this assumption [9] [10].

Artificial neural networks (ANN) have been widely used in the field of system identification, adaptive control, and statistical modelling in recent years. A neural network is composed of many non-linear adaptive processing elements and is capable of approximating any measurable function under certain conditions.

To improve the controller response of a DC-DC converter to dynamical system changes, neural network has been chosen as an alternative to classic methods ([7], [8]). However, no prior work has been done in the neural network approach for control of a PSFB (Phase-Shifted Full-Bridge) converter yet. In this research, a multi-layer feed-forward neural network controller is proposed. A Matlab Simulink model is developed first to generate training data; then the neural network is trained by Levenberg-Marquardt back-propagation algorithm. Finally, the neural controller is implemented and tested on a DSP evaluation board eZdsp® F2812 (with digital processor TMS320F2812) and a phase-shifted zero-voltage-switching circuit board UCC3895EVM. The neural network controller is fine-tuned on-line with back-propagation algorithm. Satisfactory experimental results are obtained.

II. CIRCUIT ANALYSIS

The circuit diagram of a PSFB DC-DC converter is shown in Fig. 1. Based on circuit analysis, the control scheme to drive the switching MOSFET (metal-oxide-semiconductor field-effect transistor, i.e., Q1, Q2, Q3 and Q4) is very complicated [6]. The circuit is operated in one of the following modes:

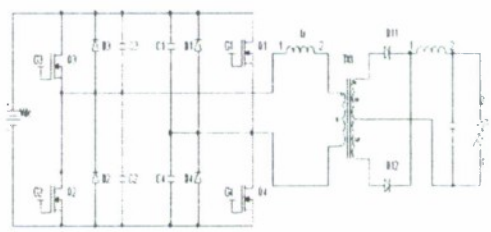


Fig. 1. PSFB DC-DC Converter

Mode 1: The diagonal MOSFET Q3 & Q4 are conducting and power is delivered through the transformer to the load. The primary load current is flowing through the leakage inductance of the transformer. The total primary current is

1-4244-0818-0/07/$20.00 (c) 2007 IEEE

equal to the load current plus increasing magnetizing current of the transformer.

Mode 2: With Q4 on, the capacitance across Q1 is charged to +V. When Q4 turns OFF the current through the transformer inductance starts to charge the drain source capacitance of Q4, while at the same time discharges the capacitance of Q1. This action continues until the body diode of Q1 turns ON to clamp the voltage across Q1 at approximately 0.7V. The current through the transformer is sustained in the upper half of the power circuit.

Mode 3: When the voltage across Q1 reaches approximately 0V, Q1 turns ON. The time that is required for the capacitance of Q4 and Q1 to reach the desired voltage is determined by the characteristics of hardware. The current in this mode circulates through the conduction channels of Q3 and Q1.

Mode 4: Q3 turns off. The transformer current now starts to charge and discharge the capacitance of Q3 and Q2 respectively. It again requires a finite amount of time for the drain voltage of Q2 to reach "0 volt" at which point Q2 is allowed to be switched on. That means non-dissipate turn ON switching is accomplished.

Mode 5: With the complete discharge of its drain source capacitance Q2 is now ready to turn ON. Power is delivered to the load through the conduction path of Q1 and Q2 for an amount of time that is determined by the control circuit. Twice the product of this time, times the operating frequency of the oscillator gives the duty cycle of the converter as in a regular PWM converter. Duty Cycle D = 2*Ton*Fs.

Mode 6: Following the power transfer by the above diagonal pair, Q1 turns OFF. The voltage across Q4 starts to decrease, and when this voltage reaches 0 volts the next mode starts.

Mode 7: In this mode, Q4 turns on and primary current circulates in the conduction channels of the lower pair.

Mode 8: Q2 turns OFF and current starts to charge and discharge the capacitance of Q2 and Q3 respectively. When the voltage across Q3 has reached 0 volts then Q3 turns ON non-dissipative and the complete cycle repeats itself from Mode 1.

The relationship between the input/output voltage and the duty cycle can be described as:

$$V_0 = \left(\frac{2nV_{in}}{T_s}\right)\left(\Phi\cdot\frac{T_s}{2} - T_1 - \frac{T_2}{2}\right) \tag{1}$$

where V_0 is the output voltage, V_{in} is the input voltage, both in rms value; n is the transformer ratio, T_s is the switching period constant, and Φ is the duty cycle. T_1 and T_2 can be obtained from the following equations:

$$T_1 = \frac{(nI + I_c)L_r}{V_{in}} \tag{2}$$

$$T_2 = \frac{V_{in}(C_{Q3} + C_{Q4})}{nI} \tag{3}$$

where I is the load current; C_{Q3} and C_{Q4} are the capacitance of the two intrinsic capacitors of Q3 and Q4, respectively, L_r is the resonant inductance. I_c can be calculated as follows:

$$I_c = V_{in}\sqrt{\frac{(C_{Q1} + C_{Q2})(J^2 - 1)}{L_r}} \tag{4}$$

where

$$J = \frac{nI}{V_{in}}\sqrt{\frac{L_r}{(C_{Q1} + C_{Q2})}} \tag{5}$$

The controller determines the duty cycle based on the input voltage and load current to achieve the desired output voltage. As shown in the above equations, the control law is highly nonlinear and as the result, it leads to a very complicated design in conventional approach.

III. LOAD CURRENT ESTIMATION

Based on the above the analysis, the Simulink model for load current estimation is developed and shown in Fig. 2, which illustrates the relationship between the change in load current and the change in the output voltage.

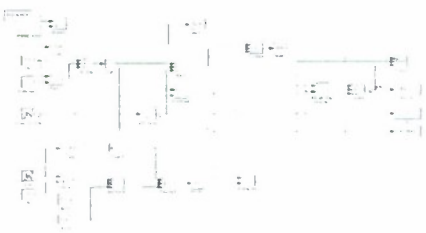


Fig. 2. Load Current Dynamic Model

It is well-known that artificial neural networks can approximate any nonlinear function to the desired accuracy, after it is fully trained. In this research, the Simulink model is used to obtain data set for the initial training of neural network. A feedforward neural network with six neurons in the hidden layer is employed. The activate function in the hidden layer is the Tangent Sigmoid function:

$$f(x) = \frac{2}{1 + e^{-2x}} - 1 \tag{6}$$

The weights of the neural network are initialized at random. To speed up the learning process, Levenberg-Marquardt algorithm is employed:

$$W(k+1) = W(k) + \Delta W \tag{7}$$

where

$$\Delta W = (J_a^T J_a + \mu I)^{-1} J_a^T e \tag{8}$$

J_a (also called the Jacobian matrix) is the first order derivative of the error function respect to the neural network weight, e is the output error (i.e., the difference between the neural network output and the desired output), μ is a learning parameter, and k is the number of iterations.

The neural network learning curve is shown in Fig. 3. From the simulation result we conclude that the neural network can successfully estimate the load current.

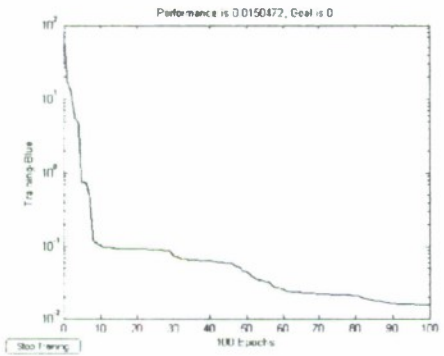


Fig. 3. Training Performance on Load Current Estimation

IV. THE NEURAL NETWORK CONTROLLER SIMULATION

Fig. 4 describes the dynamic model for overall system, with the estimation of load current, the input voltage, and the duty cycle. A feed-forward neural network with one hidden layer is employed to calculate the desired duty cycle. The inputs to the neural network include input voltage, load current, and the change of output voltage.

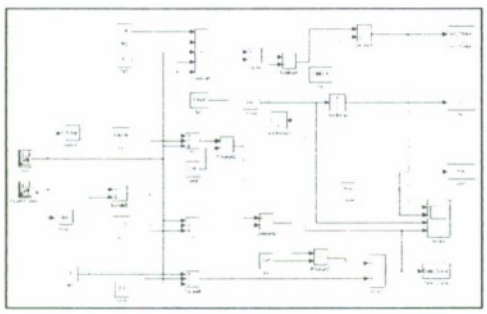


Fig. 4. The Simulink Model

There are six neurons in the hidden layer and one neuron in the output layer. The activation function for each neuron is:

$$f(x) = \frac{1}{1+e^{-x}} \qquad (9)$$

The computer simulation results are shown in the following figures. Fig. 5 shows the training error is reduced to the desired accuracy after about 100 iterations. Fig. 6 demonstrates the desired value of duty cycle and the estimation of neural network controller. It is shown that the neural network can be fully trained to model the relationship between duty cycle and the change of output voltage.

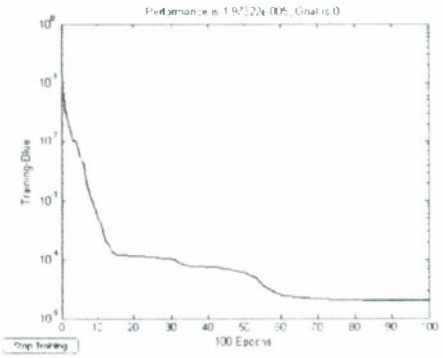


Fig. 5. Neural Network Training Error

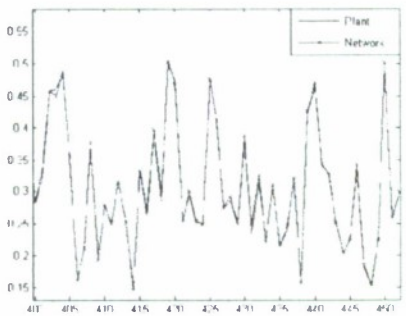


Fig. 6. Duty Cycle Estimation

V. EXPERIMENTAL RESULTS

In order to fully investigate the performance of the neural network controller, a DSP evaluation module eZdsp® F2812 and a phase-shifted zero-voltage-switching evaluation board UCC3895EVM are used to implement the neural controller and the power circuit, respectively. The on-board digital signal processor TMS320F2812 is a 32-bit CPU with an operating speed of 150 MIPS (million instructions per second). The input voltage of the phase-shifted full-bridge converter ranges from 36V to 72V, with 48V as its nominal input. The nominal value of the output of the converter is 3.3V, with nominal load current. The neural network controller is developed in C language. The overall system schematic is shown in Fig. 8 (next page).

First, the analog signals from the power circuit are digitized and processed before being fed into the neural

network's inputs. Since the power circuit is driven by 400 kHz high frequency PWM signals, the raw output voltage signal is quite noisy during the switch turn-on and turn-off time. To solve this problem, multiple samples are taken to obtain the average value over a certain period of time. Next, the processed data is fed into the neural network to estimate the desired duty cycle. The new estimated duty cycle is then applied to control the power circuit, and the output voltage is monitored.

At nominal input voltage 48V, the 10% to 90% load step up and 90% to 10% load step down transient response of the overall system is captured on the oscilloscope as shown in Fig. 7. The difference between the voltage peak and the nominal output is about 0.25 to 0.3V for both load step up and step down transient response, which is well within the normal range of switching regulator's transient performance. However, it is observed that the converter's response time to load transient is longer than the case of analog controller. This is due to the on-line training and long processing time for floating-point format running on a fixed-point microprocessor.

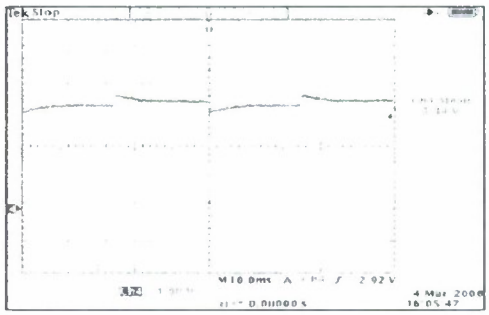


Fig. 7. 10% ←→ 90% Load Transient Response at 48V with Neural Network Controller

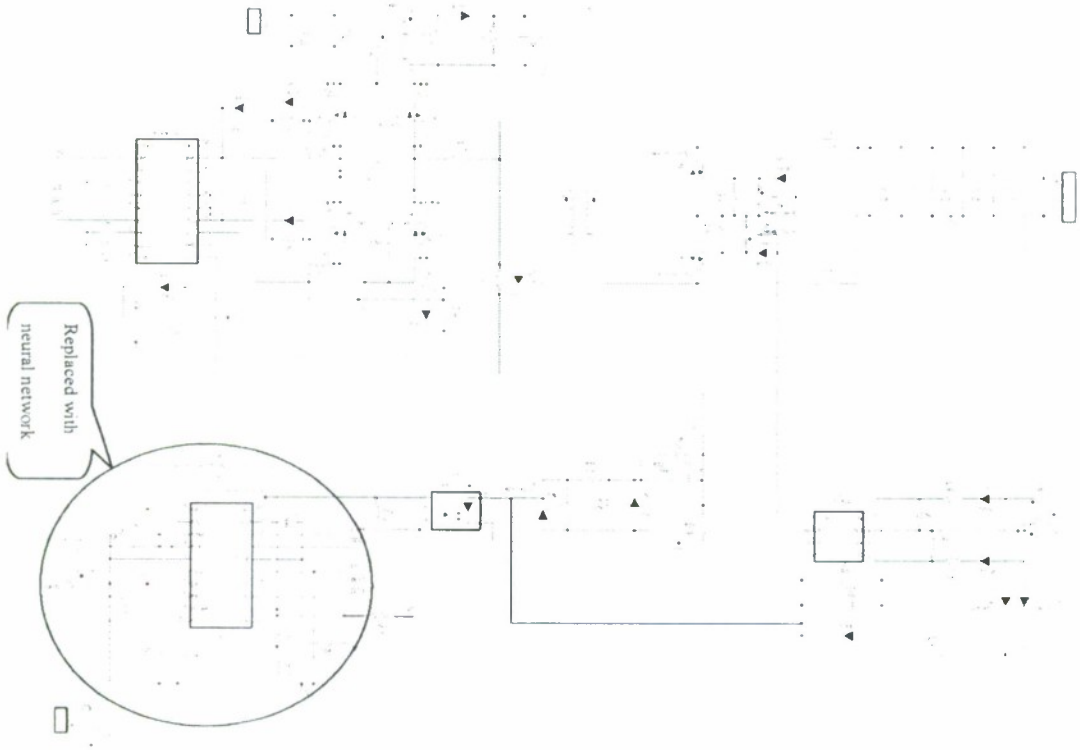


Fig. 8. Power Circuit and Controller Schematic

Another test is performed when the input voltage is 60V. The 10% to 90% load step up and 90% to 10% load step down transient response of the overall system is shown in Fig. 9. The difference between the voltage peak and the nominal output is about 0.4 to 0.5V for transient response. In both cases (48V and 60V input voltage), when the load changes, the neural network controller successfully generates the control (i. e., duty cycle) to keep the output voltage stable, with the desired level.

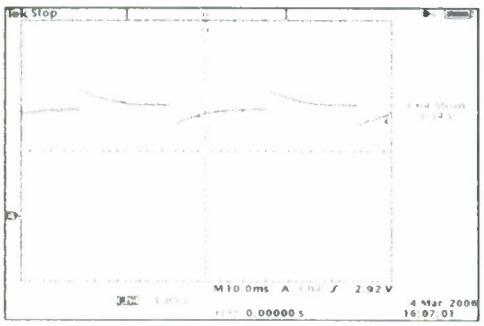


Fig. 9. 10% ←→ 90% Load Transient Response at 60V with neural network controller

It is known that phase-shifted zero-voltage switching full-bridge converters employ zero-voltage-switching (ZVS) technique which allows the voltage across each transistor to swing to zero just before the start of the next conduction cycle ([4], [5], [6]) to reduce power consumption and thus improve efficiency. This approach can also minimize the switching stress on the transistor so that the longevity of the device is extended. Fig. 10 shows the oscilloscope capture of switching waveform of MOSFET Q2 in the power circuit, where V_g is the driving signal at the gate terminal, and V_{ds} is the voltage across drain and source, respectively. The waveforms clearly show that the voltage across the MOSFET swings to zero before the driving signal is asserted. Therefore, the transistor is turned on with zero cross voltage and the switching loss is greatly reduced. Similar waveforms can be captured for other MOSFETs in the converter.

Finally, the performance of on-line training algorithm is investigated. The output voltages at varying load conditions (with nominal input voltage) are measured and compared, as shown in Fig. 11. The off-line training is based on an ideal Simulink model (without considering the voltage drop and power consumption of each circuit element); therefore the resulting output voltage is significantly lower than the desired value. With the correction of on-line training, the performance of neural network controller is significantly improved and performs as expected.

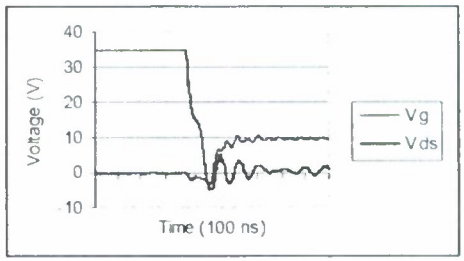


Fig. 10. Zero-Voltage Switching

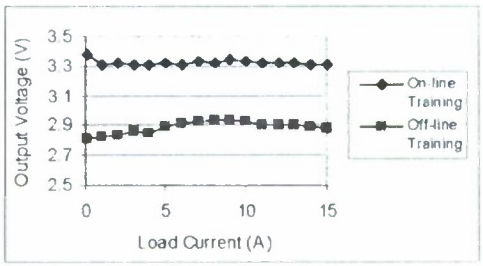


Fig. 11. On-line vs. Off-line Training

VI. CONCLUSION

A neural controller as an alternative to classic controller for PSFB DC-DC converter is proposed in this paper. Both simulation and experimental results show that the neural network controller is able to estimate the desired duty cycle under several dynamic conditions. More research works will be done to optimize the software to speed up the neural controller for real-time applications.

ACKNOWLEDGEMENTS

This work was partially sponsored by the Department of the Navy, Office of Naval Research, under Award # N00014-06-1-1111. Also, the authors would like to thank Dr. Taufik for providing the hardware.

REFERENCES

[1] P. R. Chetty, "Resonant power supplies: Their history and status," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 7, no. 4, pp. 23–29, Apr. 1992.

[2] M. G. Kim and M. J. Youn, "An energy feedback control of series resonant converters," *IEEE Trans. Power Electron.*, vol. 6, no. 4, pp. 338–345, Jul. 1991.

[3] J. M. Carrasco, E. Galván, G. E. Valderrama, R. Ortega, and A. Stankovic, "Analysis and experimentation of nonlinear adaptive controllers for the series resonant converter," *IEEE Trans. Power Electron.*, vol. 15, no. 3, pp. 536–544, May 2000.

[4] H.-S. Choi and B. H. Cho, "Novel zero-current-switching (ZCS) PWM switch cell minimizing additional conduction loss," *IEEE Trans. On Industrial Electronics*, Vol. 49, No. 1, P. 165-172, February 2002.

[5] H.-S. Choi, J.-W. Kim, and B. H. Cho, "Novel zero-voltage and zero-current-switching (ZVZCS) full-bridge PWM converter using coupled output inductor," *IEEE Trans. On Power Electronics*, Vol. 17, No. 5, P. 641-648, September 2002.

[6] X. Ruan and Y. Yan, "A novel zero-voltage and zero-current-switching PWM full-bridge converter using two diodes in series with the lagging leg," *IEEE Trans. On Industrial Electronics*, Vol. 48, No. 4, P. 777-785, August 2001.

[7] J. M. Quero, J. M. Carrasco, and L. G. Franquelo, "Implementation of a neural controller for the series resonant converter," *IEEE Trans. On Industrial Electronics*, Vol. 49, No. 3, P. 628-639, June 2002.

[8] F. Kamran, R. G. Harley, B. Burton, T. G. Habetler, and M. A. Brooke, "A fast on-line neural-network training algorithm for a rectifier regulator," *IEEE Trans. On Power Electronics*, Vol. 13, No. 2, March 1998.

[9] B. Choi, J. Kim, B. H. Cho, S. Choi, and C. M. Wildrick, "Designing control loop for DC-to-DC converters loaded with unknown AC dynamics," *IEEE Trans. On Industrial Electronics*, Vol. 49, No. 4, P. 925-932, August 2002.

[10] Texas Instrument Inc, "BiCMOS advanced phase shift PWM controller," Data sheet of UCC3895, January 2001.

Hindawi Publishing Corporation
International Journal of Rotating Machinery
Volume 2008, Article ID 846365, 11 pages
doi:10.1155/2008/846365

Research Article

A Two-Disk Extended Jeffcott Rotor Model Distinguishing a Shaft Crack from Other Rotating Asymmetries

Xi Wu and Jim Meagher

Department of Mechanical Engineering, California Polytechnic State University, San Luis Obispo, CA 93407, USA

Correspondence should be addressed to Xi Wu, xwu@calpoly.edu

Received 23 August 2007; Revised 26 November 2007; Accepted 12 January 2008

Recommended by Eric Maslen

A mathematical model of a cracked rotor and an asymmetric rotor with two disks representing a turbine and a generator is utilized to study the vibrations due to imbalance and side load. Nonlinearities typically related with a "breathing" crack are included using a Mayes steering function. Numerical simulations demonstrate how the variations of rotor parameters affect the vibration response and the effect of coupling between torsional and lateral modes. Bode, spectrum, and orbit plots are used to show the differences between the vibration signatures associated with cracked shafts versus asymmetric shafts. Results show how nonlinear lateral-torsional coupling shifts the resonance peaks in the torsional vibration response for cracked shafts and asymmetric rotors. The resonance peaks shift depending on the ratio of the lateral-to-torsional natural frequencies with the peak responses occurring at noninteger values of the lateral natural frequency. When the general nonlinear models used in this study are constrained to reduce to linear torsional vibration, the peak responses occur at commonly reported integer ratios. Full spectrum analyses of the X and Y vibrations reveal distinct vibration characteristics of both cracked and asymmetric rotors including reverse vibration components. Critical speeds and vibration orders predicted using the models presented herein include and extend diagnostic indicators commonly reported.

1. INTRODUCTION

The purpose of this investigation is to develop and test models for the vibration response of cracked and asymmetric shafts. Some asymmetries are geometric while others may be due to a shaft crack. In this paper, an asymmetric shaft refers to geometric asymmetry other than that due to a crack. The vibration response of asymmetric and cracked shafts shares characteristics such as $2x$ response which makes them hard to distinguish.A distinct crack diagnostic measure observable with measurable vibration data is a further goal of this study. This topic is widely studied because of possible sudden catastrophic failure of a rotor from fatigue. Stress concentrations and high-rotational speeds exacerbate the problem. This is especially dangerous because the torsional response of the rotor is often unmeasured and lightly damped. A comprehensive literature survey of various crack modeling techniques and system behavior of cracked rotor was given by Wauer [1]. This paper contains the modeling of the cracked components of the structures and searches for different detection strategies to diagnose fracture damage. A more recent survey paper by Sabnavis et al. [2] divides the current research into three categories: vibration-based methods, modal testing, and nontraditional methods such as wavelets or neural networks. Dimarogonas [3] provided an earlier literature review of the vibration of cracked structures and cites more than 300 papers. His review is categorized according to methods that describe local flexibility due to cracks, nonlinearities introduced into the system, and local stiffness matrix descriptions of the cracked section. The crack leads to a coupled system that can be recognized from additional harmonics in the frequency spectrum. The subharmonic resonances at approximately half and one third of the bending critical speed of the rotor are reported to be the prominent crack indicators by Gasch [4, 5] and Chan [6]. By utilizing a single parameter "hinge" crack model, Gasch provided an overview of the dynamic behavior of a simple rotor with transverse crack. He assumed weight dominance and employed a perturbation method into his analysis. Cross-coupling stiffness and dynamic response terms were not included in his analysis. Mayes model [7] is more practical for deep cracks than a hinged model. Based on Mayes modified

and two lateral displacements in inertial coordinates. The inboard disk's vibration is described by the angle $\Theta(t)$. The location of the center of mass of the outboard disk can be expressed as the following:

$$\begin{aligned} x_{cm} &= X + \varepsilon\cos(\Phi + \delta), \\ y_{cm} &= Y + \varepsilon\sin(\Phi + \delta). \end{aligned} \tag{1}$$

2.2. Equations of motion

The kinetic energy, potential energy, and dissipation function for the rotor system can, respectively, be expressed as the following:

$$\begin{aligned} T &= \frac{1}{2}I(\Omega + \dot{\varphi})^2 + \frac{1}{2}I_0(\Omega + \dot{\theta})^2 \\ &\quad + \frac{M}{2}\{\dot{X}^2 + \dot{Y}^2 - 2\varepsilon\dot{X}(\Omega + \dot{\varphi})\sin(\Phi + \delta) \\ &\quad + 2\varepsilon\dot{Y}(\Omega + \dot{\varphi})\cos(\Phi + \delta) + \varepsilon^2(\Omega + \dot{\varphi})^2\}, \end{aligned} \tag{2}$$

$$U = \frac{1}{2}(k_{11}X^2 + k_{22}Y^2) + k_{12}XY + \frac{1}{2}K_t(\varphi - \theta)^2,$$

$$D = \frac{1}{2}C\dot{X}^2 + \frac{1}{2}C\dot{Y}^2 + \frac{1}{2}C_t(\dot{\varphi} - \dot{\theta})^2.$$

Loads applied to the system include a driving torque applied to the inboard disk, $C_c(\Omega - \dot{\theta}) + K_c(\Omega t - \theta)$, and a vertical side load, P, applied to the outboard disk. The damping is modeled as lumped viscous damping at the outboard disk and lumped torsional viscous damping of the shaft. The stiffness matrix for a Jeffcott rotor with a cracked shaft in inertial coordinates, $\mathbf{K}_{Ic}$, is given by [5, 8, 9]. Details can be found in [15],

$$\begin{aligned} \mathbf{K}_{Ic} &= \begin{pmatrix} k_{11} & k_{12} \\ k_{21} & k_{22} \end{pmatrix} \\ &= \begin{pmatrix} K & 0 \\ 0 & K \end{pmatrix} - \frac{f(\Phi)K}{2}\begin{pmatrix} \Delta k_1 + \Delta k_2\cos 2\Phi & \Delta k_2 \sin 2\Phi \\ \Delta k_2 \sin 2\Phi & \Delta k_1 - \Delta k_2 \cos 2\Phi \end{pmatrix}, \end{aligned} \tag{3}$$

where

$$\Delta k_1 = \frac{\Delta k_\xi + \Delta k_\eta}{K}, \qquad \Delta k_2 = \frac{\Delta k_\xi - \Delta k_\eta}{K}, \tag{4}$$

$\Delta k_\xi, \Delta k_\eta$ are, respectively, the reduced stiffness in ξ and η directions in a rotor-fixed coordinate system, and $f(\Phi) = (1 + \cos(\Phi))/2$ is a steering function which Mayes and Davies [7] proposed to illustrate a smooth transition between the opening and closing of a "breathing" crack in rotating coordinates; and $\Delta k_\eta = \Delta k_\xi/6$ is assumed to describe the stiffness variation for deep cracks.

The stiffness matrix for a rotor with an asymmetric shaft in inertial coordinates is given by

$$\mathbf{K}_{Iasym} = \mathbf{T}\mathbf{K}_R\mathbf{T}^{-1}, \tag{5}$$

where $\mathbf{T}$ is the coordinate transformation matrix, $\mathbf{T} = \begin{pmatrix} \cos\Phi & -\sin\Phi \\ \sin\Phi & \cos\Phi \end{pmatrix}$ and $\mathbf{K}_R = \begin{pmatrix} K_\xi & 0 \\ 0 & K_\eta \end{pmatrix}$ is the stiffness matrix in rotating coordinates:

$$\mathbf{K}_{Iasym} = \begin{pmatrix} K + \Delta K\cos(2\Phi) & \Delta K\sin(2\Phi) \\ \Delta K\sin(2\Phi) & K - \Delta K\cos(2\Phi) \end{pmatrix}. \tag{6}$$

Note that for the asymmetric shaft, the stiffness parameters differ from the parameters used for a cracked shaft. In the asymmetric shaft model, K is the average stiffness rather than the uncracked shaft stiffness, and ΔK and q are asymmetric stiffness factors:

$$K = \frac{K_\xi + K_\eta}{2}, \qquad \Delta K = \frac{K_\xi - K_\eta}{2}, \qquad q = \frac{K_\eta - K_\xi}{2K}, \tag{7}$$

with these factors

$$\mathbf{K}_{Iasym} = \begin{pmatrix} k_{11} & k_{12} \\ k_{21} & k_{22} \end{pmatrix} = K\begin{bmatrix} 1 - q\cos 2\Phi & -q\sin 2\Phi \\ -q\sin 2\Phi & 1 + q\cos 2\Phi \end{bmatrix}. \tag{8}$$

2.3. Cracked shaft equations of motion

The general equations of motion are obtained using Lagrange's equations. For the cracked shaft, the equations of motion become

$$\begin{aligned} &\ddot{X} + \frac{C}{M}\dot{X} + \omega_n^2\left[1 - \frac{f(\Phi)}{2}(\Delta k_1 + \Delta k_2\cos 2\Phi)\right]X \\ &\quad - \frac{\omega_n^2 f(\Phi)\Delta k_2 \sin 2\Phi}{2}\left(Y_m - \frac{P}{K}\right) \\ &\quad = \varepsilon(\Omega + \dot{\varphi})^2\cos(\Phi + \delta) + \varepsilon\ddot{\varphi}\sin(\Phi + \delta), \\ &\ddot{Y}_m + \frac{C}{M}\dot{Y}_m - \frac{\omega_n^2 f(\Phi)\Delta k_2 \sin 2\Phi}{2}X \\ &\quad + \omega_n^2\left[1 - \frac{f(\Phi)}{2}(\Delta k_1 - \Delta k_2\cos 2\Phi)\right]Y_m \\ &\quad = \varepsilon(\Omega + \dot{\varphi})^2\sin(\Phi + \delta) - \varepsilon\ddot{\varphi}\cos(\Phi + \delta) \\ &\quad - \frac{Pf(\Phi)}{2M}(\Delta k_1 - \Delta k_2\cos 2\Phi), \\ &\ddot{\theta} + \frac{K_t + K_c}{I_0}\theta - \frac{K_t}{I_0}\varphi = -\frac{C_t + C_c}{I_0}\dot{\theta} + \frac{C_t}{I_0}\dot{\varphi}, \\ &\ddot{\varphi} + \frac{C_t}{I}\dot{\varphi} - \frac{C_t}{I}\dot{\theta} + \frac{K_t}{I}\varphi - \frac{K_t}{I}\theta \\ &\quad = \frac{P\varepsilon f(\Phi)}{2I}(\Delta k_1\cos(\Phi + \delta) - \Delta k_2\cos(\Phi - \delta)) \\ &\quad + \frac{P^2}{2KI}\left[\frac{1}{2}\frac{\partial f(\Phi)}{\partial\Phi}(\Delta k_1 - \Delta k_2\cos 2\Phi)\right. \\ &\quad \left. + f(\Phi)\Delta k_2\sin 2\Phi\right] + \Gamma_c, \end{aligned} \tag{9}$$

where

$$\begin{aligned}\Gamma_c = &-\frac{C\varepsilon}{\rho^2 M}[\ddot{X}\sin(\Phi+\delta) - \ddot{Y}_m\cos(\Phi+\delta)] \\ &+ \frac{\varepsilon\omega_n^2}{\rho^2}[-X\sin(\Phi+\delta) + Y_m\cos(\Phi+\delta)] \\ &+ \frac{\varepsilon q\omega_n^2}{\rho^2}[-X\sin(\Phi-\delta) + Y_m\cos(\Phi-\delta)] - \frac{qP}{I}Y_m\sin 2\Phi \\ &+ \frac{q\omega_n^2}{\rho^2}\left[\left(Y_m\left(Y_m - \frac{P}{K}\right) - X^2\right)\sin 2\Phi \right. \\ &\left. + 2X\left(Y_m - \frac{P}{K}\right)\cos 2\Phi\right].\end{aligned} \tag{15}$$

Using nondimensionalized time defined by (11), (14), and (15) takes the following form:

$$\begin{aligned}&X'' + 2\zeta X' + (1 - q\cos 2\Phi)X - q\left(Y_m - \frac{P}{M\omega_n^2}\right)\sin 2\Phi \\ &\quad = \varepsilon\left(\frac{\Omega}{\omega_n} + \varphi'\right)^2\cos(\Phi+\delta) + \varepsilon\varphi''\sin(\Phi+\delta), \\ &Y_m'' + 2\zeta Y_m' - qX\sin 2\Phi + (1 + q\cos 2\Phi)Y_m \\ &\quad = \varepsilon\left(\frac{\Omega}{\omega_n} + \varphi'\right)^2\sin(\Phi+\delta) - \varepsilon\varphi''\cos(\Phi+\delta) \\ &\qquad + \frac{P}{M}\frac{q}{\omega_n^2}\cos 2\Phi, \\ &\theta'' + R_I(1+K_r)\left(\frac{\omega_t}{\omega_n}\right)^2\theta - R_I\left(\frac{\omega_t}{\omega_n}\right)^2\varphi \\ &\quad = -2R_I\zeta_t(1+C_r)\frac{\omega_t}{\omega_n}\theta' + 2R_I\zeta_t\frac{\omega_t}{\omega_n}\varphi', \\ &\varphi'' + 2\zeta_t\frac{\omega_t}{\omega_n}\varphi' - 2\zeta_t\frac{\omega_t}{\omega_n}\theta' + \left(\frac{\omega_t}{\omega_n}\right)^2\varphi - \left(\frac{\omega_t}{\omega_n}\right)^2\theta \\ &\quad = \left(\frac{P}{M}\right)^2\frac{q}{\omega_n^4\rho^2}\sin 2\Phi \\ &\qquad - \frac{P}{M}\frac{q\varepsilon}{\omega_n^2\rho^2}\cos(\Phi-\delta) + \frac{\Gamma_c}{\omega_n^2},\end{aligned} \tag{16}$$

$$\begin{aligned}\frac{\Gamma_c}{\omega_n^2} = &-2\zeta\frac{\varepsilon}{\rho^2}[X'\sin(\Phi+\delta) - Y_m'\cos(\Phi+\delta)] \\ &+ \frac{\varepsilon}{\rho^2}[-X\sin(\Phi+\delta) + Y_m\cos(\Phi+\delta)] \\ &+ \frac{q\varepsilon}{\rho^2}[-X\sin(\Phi-\delta) + Y_m\cos(\Phi-\delta)] \\ &- \frac{Pq Y_m}{M\rho^2\omega_n^2}\sin 2\Phi + \frac{q}{\rho^2}\left[Y_m\left(Y_m - \frac{P}{M\omega_n^2}\right) - X^2\right]\sin 2\Phi \\ &+ \frac{2qX}{\rho^2}\left(Y_m - \frac{P}{M\omega_n^2}\right)\cos 2\Phi.\end{aligned} \tag{17}$$

2.5. Special cases

Case 1 (pure torsional vibration for a cracked rotor). Assuming no lateral vibration, $X = 0, Y_m = 0$, and a rigid drive coupling, $\dot{\Theta} = \Omega$, leads to the following simplification for the cracked shaft:

$$\begin{aligned}&\ddot{\varphi} + \frac{C_t}{I}\dot{\varphi} + \frac{K_t}{I}\varphi \\ &\quad = \frac{P\varepsilon}{8I}(\Delta k_1 - \Delta k_2) \\ &\qquad + \left\{\frac{P\varepsilon}{4I}(\Delta k_1 - \Delta k_2)\cos\Phi + \frac{P^2}{8KI}\left(-\Delta k_1 + \frac{\Delta k_2}{2}\right)\sin\Phi\right\} \\ &\qquad + \left\{\frac{P\varepsilon}{8I}(\Delta k_1 - \Delta k_2)\cos 2\Phi + \frac{P^2\Delta k_2}{4KI}\sin 2\Phi\right\} \\ &\qquad + \frac{3P^2\Delta k_2}{16KI}\sin 3\Phi.\end{aligned} \tag{18}$$

We introduce the following two constants:

$$E_1 = \frac{P\varepsilon}{2I\omega_n^2} = \frac{P}{2M}\frac{\varepsilon}{\rho^2\omega_n^2}, \qquad E_2 = \frac{P^2}{2KI\omega_n^2} = \frac{P^2}{2M^2}\frac{1}{\rho^2\omega_n^4}. \tag{19}$$

Using nondimensional time defined by (11) and (18) takes the following form:

$$\begin{aligned}&\varphi'' + 2\zeta_t\frac{\omega_t}{\omega_n}\varphi' + \left(\frac{\omega_t}{\omega_n}\right)^2\varphi \\ &\quad = \frac{E_1}{4}(\Delta k_1 - \Delta k_2) \\ &\qquad + \left\{\frac{E_1}{2}(\Delta k_1 - \Delta k_2)\cos\Phi + \frac{E_2}{4}\left(-\Delta k_1 + \frac{\Delta k_2}{2}\right)\sin\Phi\right\} \\ &\qquad + \left\{\frac{E_1}{4}(\Delta k_1 - \Delta k_2)\cos 2\Phi + \frac{E_2\Delta k_2}{2}\sin 2\Phi\right\} \\ &\qquad + \frac{3E_2\Delta k_2}{8}\sin 3\Phi.\end{aligned} \tag{20}$$

Case 2 (pure torsional vibration for an asymmetric rotor). Assuming no lateral vibration, $X = 0, Y_m = 0$, and a rigid drive coupling, $\dot{\Theta} = \Omega$, leads to the following simplification for the asymmetric shaft:

$$\ddot{\varphi} + \frac{C_t}{I}\dot{\varphi} + \frac{K_t}{I}\varphi = -\frac{qP\varepsilon}{I}\cos(\Phi-\delta) + \frac{qP^2}{KI}\sin 2\Phi. \tag{21}$$

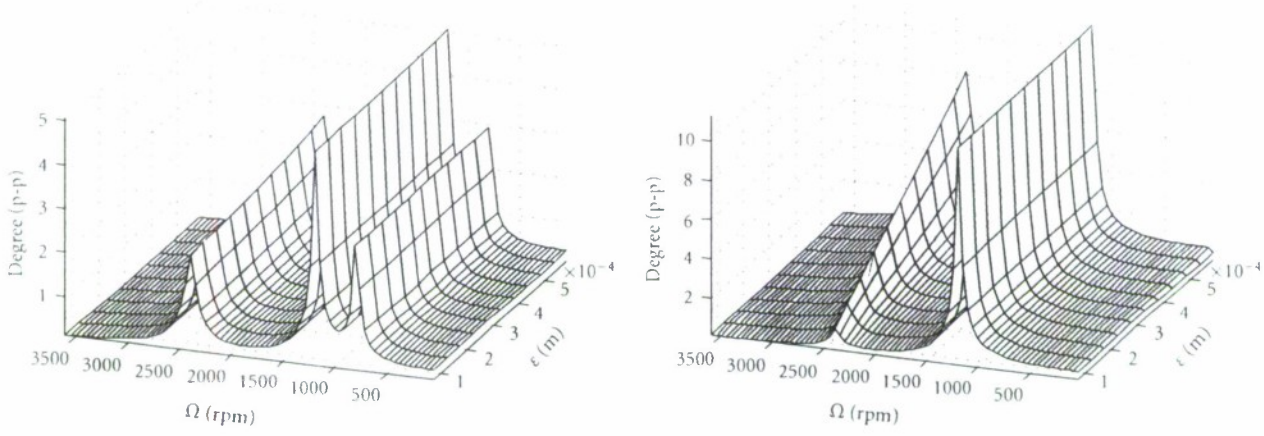
Using nondimensional time defined by (11) and (21) takes the following form:

$$\begin{aligned}&\varphi'' + 2\zeta_t\left(\frac{\omega_t}{\omega_n}\right)\varphi' + \left(\frac{\omega_t}{\omega_n}\right)^2\varphi \\ &\quad = -\frac{P}{M}\frac{q\varepsilon}{\omega_n^2\rho^2}\cos(\Phi-\delta) + \left(\frac{P}{M}\right)^2\frac{q}{\omega_n^4\rho^2}\sin 2\Phi.\end{aligned} \tag{22}$$

The equations above were solved using a variable time-step integration algorithm after the following normalization and simplifications, $Y = Y_m - P/K$ is used to delineate the static offset from dynamic response.

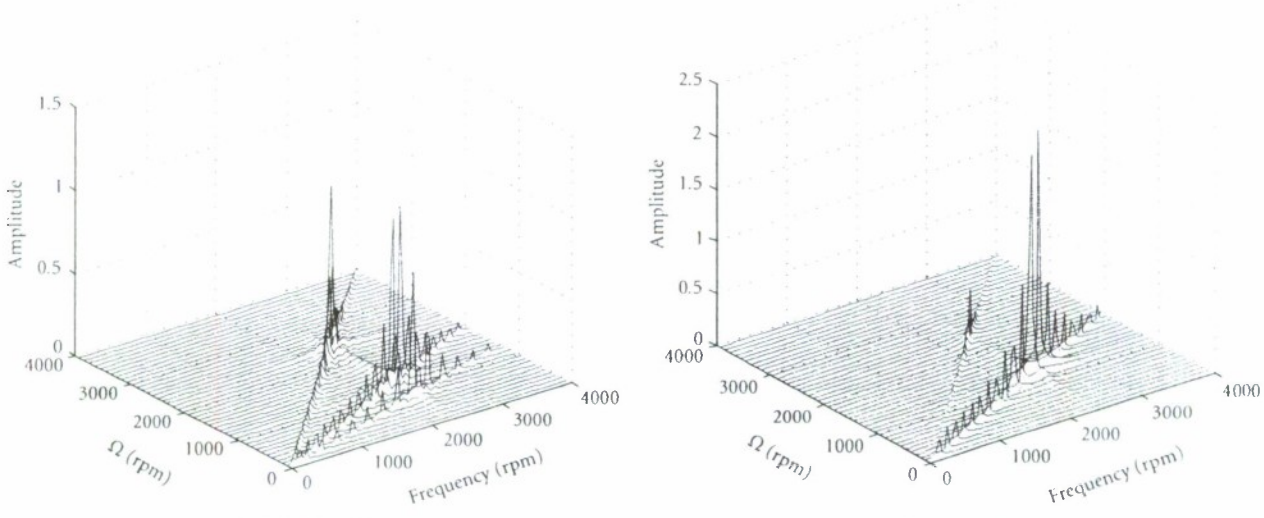
Table 1: Model physical parameters for pure torsional vibration, special cases.

Cracked shaft			Asymmetric shaft		
Parameters	Values	Units	Parameters	Values	Units
ω_n	12000	rpm	ω_n	12000	rpm
ω_t	2400	rpm	ω_t	2400	rpm
ε	7.62×10^{-5}	m	ε	7.62×10^{-5}	m
ρ	0.0241	m	ρ	0.0241	m
P/M	1270	m/s^2	P/M	1270	m/s^2
ζ_t	0.02		ζ_t	0.02	
$\Delta k_\xi/K$	0.30		q	0.14	
δ	0	rad	δ	0	rad



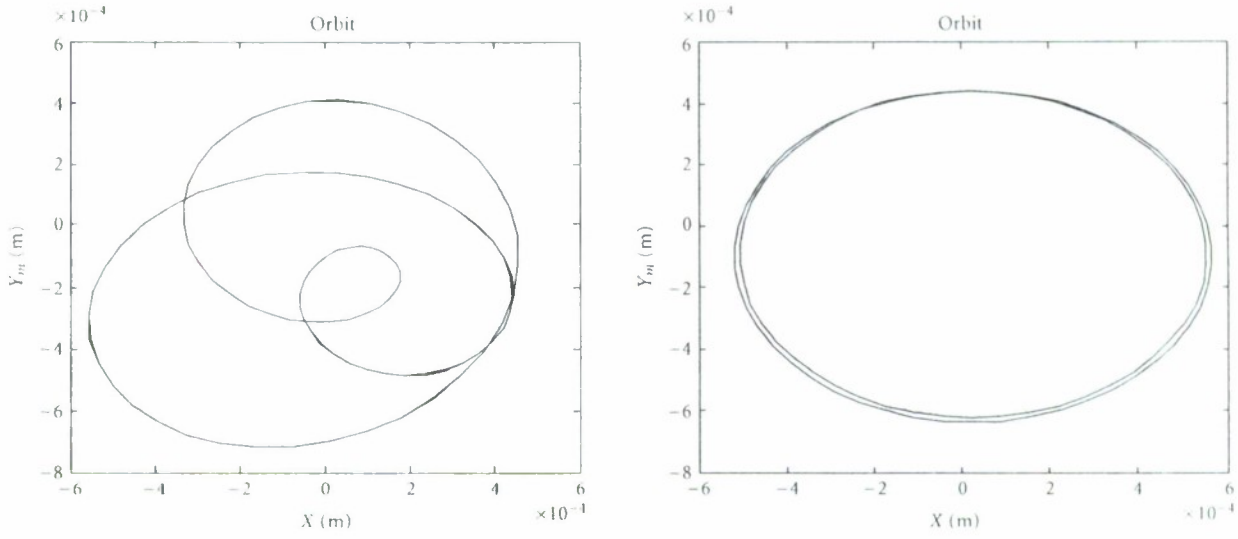
(a) $\Delta k_\xi/K = 0.30$ (b) $q = 0.14$

Figure 4: Overall peak-to-peak torsional vibration response, φ, special cases, for cracked shaft (a) and asymmetric shaft (b).



(a) $\Delta k_\xi/K = 0.30$ (b) $q = 0.14$

Figure 5: Half spectrum of zero-to-peak torsional vibration response, φ, special cases, for cracked shaft (a) and asymmetric shaft (b).



(a) Cracked shaft (b) Asymmetric shaft

FIGURE 8: Shaft whirling at $\Omega = \omega_n/3$; $\omega_t/\omega_n = 1.38$; $\omega_n = 2400$ rpm, Table 3 parameters.

TABLE 3: Model physical parameters for torsional and lateral vibration, general case.

Cracked shaft			Asymmetric shaft		
Parameters	Values	Units	Parameters	Values	Units
ω_n	2400	rpm	ω_n	2400	rpm
ω_t	3312	rpm	ω_t	3312	rpm
ε	5.08×10^{-5}	m	ε	5.08×10^{-5}	m
ρ	0.0229	m	ρ	0.0229	m
P/M	101.6	m/s^2	P/M	101.6	m/s^2
ζ_t	0.02		ζ_t	0.02	
ζ	0.1		ζ	0.1	
K_r	5		K_r	5	
C_r	1		C_r	1	
R_I	1		R_I	1	
$\Delta k_t/K$	0.30		q	0.18	
δ	0	rad	δ	0	rad

addition to $1x$ and the supersynchronous responses at $2x$ and $3x$. The asymmetric lateral shaft response contains only $1x$ and $2x$ without any reverse vibration components.

4. CONCLUSIONS

This paper documents the effect of a shaft crack versus other geometric asymmetries on lateral and torsional vibrations of a two-mass rotor system. Nondimensional analytical models of extended Jeffcott rotors are derived from Lagrange's equations taking into consideration the lateral/torsional vibration coupling mechanism induced by a "breathing" crack or a geometric asymmetry. Four degrees of freedom describe the models; two lateral displacements, one torsional angular displacement of an outboard disk, and the torsional angular displacement of an inboard disk. The nonlinearities associated with a breathing crack or geometric asymmetry couple the four equations of motion. Two cases are considered in this work: a torsionally rigid rotor without lateral vibration and a general unconstrained solution to the four degrees of freedom model presented. The first case is characterized by torsional vibrations which occur at $\Omega = \omega_t/3$ and $\Omega = \omega_t/2$. For a cracked shaft, a $3x$ torsional vibration also occurs. The general case makes evident the existence of strong coupling between lateral and torsional vibrations where vibration amplitude increases with crack depth, stiffness asymmetry, and radial load. Nonlinear lateral-torsional coupling from a crack shifts the resonance peaks in the torsional vibration response. The resonance peak frequencies shift depending on the ratio of the lateral to torsional natural frequencies with the peak

[2] G. Sabnavis, R. G. Kirk, M. Kasarda, and D. Quinn, "Cracked shaft detection and diagnostics: a literature review," *The Shock and Vibration Digest*, vol. 36, no. 4, pp. 287–296, 2004.

[3] A. D. Dimarogonas, "Vibration of cracked structures: a state of the art review," *Engineering Fracture Mechanics*, vol. 55, no. 5, pp. 831–857, 1996.

[4] R. Gasch, "Dynamic behavior of a simple rotor with a cross-sectional crack," in *Proceedings of the International Conference on Vibrations in Rotating Machinery (IMechE '76)*, pp. 123–128, Cambridge, UK, September 1976, Paper C178/76.

[5] R. Gasch, "Survey of the dynamic behaviour of a simple rotating shaft with a transverse crack," *Journal of Sound and Vibration*, vol. 160, no. 2, pp. 313–332, 1993.

[6] R. K. C. Chan and T. C. Lai, "Digital simulation of a rotating shaft with a transverse crack," *Applied Mathematical Modelling*, vol. 19, no. 7, pp. 411–420, 1995.

[7] I. W. Mayes and W. G. R. Davies, "Analysis of the response of a multi-rotor-bearing system containing a transverse crack in a rotor," *Journal of Vibration, Acoustics, Stress, and Reliability in Design*, vol. 106, no. 1, pp. 139–145, 1984.

[8] J. T. Sawicki, X. Wu, G. Y. Baaklini, and A. L. Gyekenyesi, "Vibration-based crack diagnosis in rotating shafts during acceleration through resonance," in *Nondestructive Evaluation and Health Monitoring of Aerospace Materials and Composites II*, vol. 5046 of *Proceedings of SPIE*, pp. 1–10, San Diego, Calif, USA, March 2003.

[9] J. T. Sawicki, D. E. Bently, X. Wu, G. Y. Baaklini, and M. I. Friswell, "Dynamic behavior of cracked flexible rotor subjected to constant driving torque," in *Proceedings of the 2nd International Symposium on Stability Control of Rotating Machinery (ISCORMA '03)*, pp. 231–241, Gdańsk, Poland, August 2003.

[10] J. T. Sawicki, X. Wu, A. L. Gyekenyesi, and G. Y. Baaklini, "Application of nonlinear dynamics tools for diagnosis of cracked rotor vibration signatures," in *Nondestructive Evaluation and Health Monitoring of Aerospace Materials, Composites, and Civil Infrastructure IV*, vol. 5767 of *Proceedings of SPIE*, pp. 286–297, San Diego, Calif, USA, March 2005.

[11] G. Mani, D. D. Quinn, and M. Kasarda, "Active health monitoring in a rotating cracked shaft using active magnetic bearings as force actuators," *Journal of Sound and Vibration*, vol. 294, no. 3, pp. 454–465, 2006.

[12] A. K. Darpe, K. Gupta, and A. Chawla, "Coupled bending, longitudinal and torsional vibrations of a cracked rotor," *Journal of Sound and Vibration*, vol. 269, no. 1-2, pp. 33–60, 2004.

[13] A. Muszynska, P. Goldman, and D. E. Bently, "Torsional/lateral vibration cross-coupled responses due to shaft anisotropy: a new tool in shaft crack detection," in *Proceedings of the International Conference on Vibrations in Rotating Machinery (IMechE '92)*, pp. 257–262, Bath, UK, 1992, Paper C432-090.

[14] D. E. Bently, P. Goldman, and A. Muszynska, ""Snapping" torsional response of an anisotropic radially loaded rotor," *Journal of Engineering for Gas Turbines and Power*, vol. 119, no. 2, pp. 397–403, 1997.

[15] X. Wu, "Vibration-based crack-induced damage detection of shaft-disk system," Doctoral dissertation, Cleveland State University, Cleveland, Ohio, USA, 2005.

[16] X. Wu, J. Meagher, and C. Judd, "Investigation of coupled lateral and torsional vibrations of a cracked rotor under radial load," in *Proceedings of the 25th International Modal Analysis Conference (IMAC '07)*, Society for Experimental Mechanics, Orlando, Fla, USA, February 2007.

Abstract Submitted
for the MAR07 Meeting of
The American Physical Society

Sorting Category: 07.2 (E)

Structural comparison of Ag-Ge-S bulk glasses and thin films FEI WANG, MUKUL JAIN, PORTER DUNN, CARTER DE LEO, EE Department, California Polytechnic State University, PUNIT BOOLCHAND, Department of ECECS, University of Cincinnati — Ternary glasses of composition $(GeS_3)_{1-x}Ag_x$ (x=0.1 and 0.2) are studied in form of bulk and thin films. Bulk glasses are synthesized and examined in Raman scattering and SEM. Raman scattering results of bulk glasses show that with increasing x, an increasing fraction of the Ag additive enters the base glass as Ag^+ with S^- anions serving to form thiogermanate species with one, two and three non-bridging S^- species. SEM measurements of the bulk glass show the material is intrinsically phase separated. White colored islands are observed distributed in a dark base. The EDS measurements show islands are Ag rich and the base is relatively Ag deficient. The Ag rich islands are expected to be mainly glassy phase Ag_2S. Thin films of same compositions are fabricated using thermal evaporation. Films are evaporated following two different procedures to prevent the material from spitting. One method was preheating outgas and the other method was using tungsten mesh wrapped boats. The stoichiometry and molecular structure of films under each procedure are analyzed by Raman scattering and SEM to be compared with bulk glasses.

Fei Wang
fwang@calpoly.edu
EE Department, California Polytechnic State University

☐ Prefer Oral Session
☒ Prefer Poster Session

Date submitted: 12 Dec 2006 Electronic form version 1.4

The California Central Coast Research Partnership: Building Relationships, Partnerships, and Paradigms for University-Industry Research Collaboration

FINAL REPORT
ONR GRANT NO. N00014-06-1-1111
September 30, 2006 to December 31, 2008

Principal Investigator:

Susan Opava, Ph.D.
Dean of Research and Graduate Programs
California Polytechnic State University
San Luis Obispo, CA

March 30, 2009