

Applied Information Systems Research Program

(AISRP)

Workshop III

Meeting Proceedings

Laboratory for Atmospheric and Space Physics
University of Colorado
Boulder, Colorado

August 3-6, 1993

N95-34218
--THRU--
N95-34243
Unclas

(NASA-CR-199103) APPLIED
INFORMATION SYSTEMS RESEARCH
PROGRAM (AISRP) WORKSHOP 3 MEETING
PROCEEDINGS Final Report (Colorado
Univ.) 235 p

G3/82 0062760

Proceedings Issued By:

Information Systems Branch
Office of Space Science
NASA Headquarters



APPLIED INFORMATION SYSTEMS RESEARCH PROGRAM

WORKSHOP III FINAL REPORT

TABLE OF CONTENTS

Executive Summary	5
Opening Remarks [Bredekamp]	7
An Overview of the Software Support Laboratory [Davis]	8
ARTIFICIAL INTELLIGENCE TECHNIQUES SESSION	
Holographic Neural Networks for Space and Commercial Applications [Liu]	9
The SEIDAM (System of Experts for Intelligent Data Management) Project [Goodenough]	10
SIGMA: An Intelligent Model-Building Assistant [Keller]	10
Multivariate Statistical Analysis Software Technologies for Astrophysical Research Involving Large Databases [Djorgovski]	11
SCIENTIFIC VISUALIZATION SESSION	
The Grid Analysis and Display System (GrADS): A Practical Tool for Earth Science Visualization [Kinter]	13
A Distributed Analysis and Visualization System for Model and Observational Data [Wilhelmson]	13
Handling Intellectual Property at UCAR/NCAR [Buzbee]	14
NCAR Graphics - What's New in 3.2? [Lackman]	15
McIDAS-eXplorer: A Version of McIDAS-X for Planetary Applications [Limaye]	15
Experimenter's Laboratory for Visualized Interactive Science (ELVIS) [Hansen]	16
SAVS: A Space and Atmospheric Visualization Science System [Szuszczewicz]	17

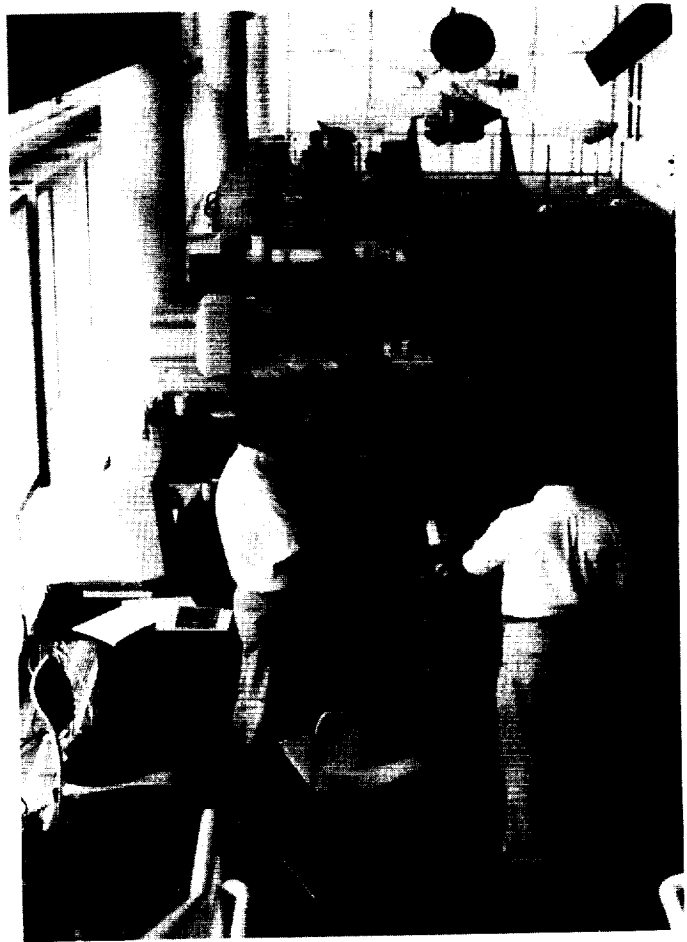
TABLE OF CONTENTS (continued)

LinkWinds - The Linked Windows Interactive Data System [Walton]	18
DataHub [Handley]	18
Data Visualization and Sensor Fusion [McCullough]	19
DATA COMPRESSION/ARCHIVING/ACCESS/ANALYSIS SESSION	
Data Compression [Storer]	20
SAMS: A Spatial Analysis and Modeling System for Environmental Monitoring [Stetina]	21
Land Surface Testbed for EOSDIS [Kelly]	21
Development of a Tool-Set for Simultaneous, Multi-Site Observations of Astronomical Objects [Chakrabarti]	22
Geographic Information System for Fusion and Analysis of High Resolution Remote Sensing and Ground Truth Data - Progress Report [Freeman]	23
Envision: A System for Management and Display of Large Data Sets [Bowman]	24
RESEARCH AND TECHNOLOGY SESSION	
Research and Technology Activities at JPL [Walton]	24
Navigation Ancillary Information Facility - An Overview of SPICE [Walton]	25
Using the NAIF SPICE Kernel Concepts and the NAIF Toolkit Software for Geometry Parameter Generation and Observation Visualization [Simmons]	26
EOS/Pathfinder Interuse Experiment [Botts]	27
Overview of Ames Research Center Advanced Network Applications [Yin]	27
Summary and Action Items [Mucklow]	28

APPENDICES

- A. AISRP Agenda
- B. Participants and Attendees
- C. Abstracts
- D. Demonstrations
- E. Presentation Material

Right:
Workstations provided attendees with first-hand demonstrations of the results of research projects.



Below:
Through demonstrations, new visualization techniques were explained to Jim Dodge and other workshop participants.



EXECUTIVE SUMMARY

The third Workshop of the Applied Information Systems Research Program (AISRP) was again hosted by the University of Colorado's Laboratory for Atmospheric and Space Physics (LASP) in Boulder, Colorado, August 3-6, 1993. The Workshop was sponsored and chaired by Glenn Mucklow of the Information Systems Branch (ISB) of NASA's Office of Space Science. The focus of this year's Workshop was "Technology Transfer," and focussed on the "applied" aspect of the Applied Information Systems Research Program. The AISRP investigators presented their progress to date, gave demonstrations of their tools and software, and addressed the technology transfer activities that have occurred and are planned or anticipated.

The presentations were organized into four sessions: Artificial Intelligence Techniques, chaired by Dr. Richard Keller from Ames Research Center; Scientific Visualization, chaired by Dr. Theo Pavlidis from State University of New York; Data Management and Archiving, chaired by Dr. James Storer from Brandeis University; and Research and Technology, chaired by Dr. Amy Walton from the Jet Propulsion Laboratory. Mr. Joseph Bredekamp, Head of the ISB, opened the Workshop with some remarks on the recent NASA reorganization and new directions for the agency.

Mr. Bredekamp encouraged the group to look at the topic of technology transfer, and discuss some lessons learned on teaming and partnerships. One of the evaluation criteria in future NASA Announcements of Opportunity will be technology transfer. In the past, there has been policy conflict surrounding this issue--the administration wants to promote U.S. competitiveness, and the science focus is on cooperative agreements and collaboration. The program is working on ways to promote both.

Before the sessions started, Dr. Randy Davis, our host from LASP, described the Software Support Laboratory (SSL), established as a result of last year's Workshop, which provides researchers with a "one-stop" location to learn about the tools that are available and how they can be accessed or obtained. It also hopes to be a conduit of information to software developers as well as provide the science community with products and trends.

During the general discussion sessions, several topics of continuing interest to the participants were raised: issues related to technology transfer, including how to better get the new tools out into the research community; issues related to hardware portability and design of tools and products; meta data and data description, including a standard framework in which to provide data information; issues related to archiving data, including the role of government and the private sector, as well as market needs as drivers; and network accessibility for researchers.

One of the challenges of a successful tool is support for users. In the current environment of decreasing financial resources, tools must be designed so that they are easy to use and maintain. The traditional software maintenance approach is too expensive for both NASA and the community. The collaboration between tool developers and tools users, which is emphasized in the AISRP, will help to meet this challenge.



Glenn H. Mucklow
Program Manager
Information Systems Research and Technology

Tuesday, August 3, 1993

Mr. Glenn Mucklow, the Workshop Chairman from NASA's Office of Space Science, welcomed the participants and thanked the Workshop host, Dr. Randal Davis and the Laboratory for Atmospheric and Space Physics (LASP) at the University of Colorado. The theme for this year's Workshop was "Technology Transfer," and focussed on the "applied" aspect of the Applied Information Systems Research Program (AISRP).

Opening Remarks

Mr. Joe Bredekamp

Information Systems Branch (ISB)

NASA Headquarters Office of Space Science

The AISRP's progress has been very encouraging, and it is developing products and starting to see some payback. In addition, the program is having a broadening influence within NASA and in the science community.

The new NASA Administrator is an agent of change, reflecting the direction of the new administration. A major message that is consciously being conveyed both internally and externally is that NASA must change its way of doing business. There have been major reorganizations within the agency. The Office of Space Science and Applications (OSSA) has been reorganized into three science program offices: Mission To Planet Earth (MTPE), the Office of Space Science (OSS), and the Office of Life and Microgravity Sciences and Applications (OLMSA). The Information Systems Branch (ISB), which supports all three science program offices and resides organizationally in the Office of Space Science, was kept as whole, integrated, and coherent as possible. The ISB has been going through a very positive transition over the past several months, but it has become clear that no assumptions can be made about what is securely in the program, and the organization will be looking carefully at everything. The administration is focussing on a different strategy--exploiting technology for broader applications. The ISB has been developing an integrated technology strategy that is aimed at developing, utilizing, and transferring technology to improve mission effectiveness and infusing this technology into the mainstream of the commercial marketplace. Information technology has an opportunity to make a significant difference in science operations and mission operations to reduce costs and increase the science return.

OSS's Process Assessment Team (PAT), chaired by Mary Kicza, is currently working in partnership with the Office of Advanced Concepts and Technology (OACT) to come up with a plan to identify technologies, infuse them into missions, and transfer them to the marketplace. The organization is also working with the Office of Space Communications to exploit ground operations, and will be doing flight/ground trade studies to try to reduce costs and improve effectiveness. Mr. Bredekamp indicated that the final draft of the Integrated Technology Strategy is in preparation, and that copies would be distributed to the AISRP group.

Mr. Bredekamp encouraged the group to look at the topic of technology transfer, and discuss some lessons learned on teaming and partnerships. One of the evaluation criteria in future AO's will be technology transfer. In the past, there has been policy conflict surrounding this issue--the administration wants to promote U.S. competitiveness, and the science focus is on cooperative agreements and collaboration. The program is working on ways to promote both.

An Overview of the Software Support Laboratory

Dr. Randal Davis

LASP, University of Colorado

One outcome from last year's Workshop was the establishment of a Software Support Laboratory at LASP to provide researchers a "one-stop" location to learn about the tools that are available and how they can be accessed or obtained.

The SSL is both a research project and a working software repository and distribution center for the AISRP, and hopes to be a conduit of information to software developers as well as provide the science community with products and trends.

Various levels of data management entities have grown up to provide better services for distributing data to the science community (through discipline data management units), and have begun to develop standards and practices within disciplines. The same type of thing needs to be done for software. Although the NASA archive, COSMIC, attempts to do this to some extent, it does not have the resources to support software and keep it active and viable for the science community.

After last year's meeting, this issue was addressed, and three models were identified regarding software distribution: 1) another party maintains and distributes software products and documentation; 2) the SSL maintains and distributes software and documentation (either via minimal or full support); and 3) the party developing or maintaining the software is seamlessly linked with the SSL's information and order services.

One of the common questions is what happens when industry steps in as partners and wants to modify the software. The issue arises when the modified product is of commercial value and there is desire to put software on ftp. It is unclear what role the SSL should have in commercialization. At the present, the primary purpose of the SSL is to facilitate a smooth transition on developed products and remain useful to the science community.

The SSL has on-line information services for software developers and users--the World Wide Web (WWW) and the Wide Area Information Service (WAIS). Dr. Davis gave a demonstration of the on-line service. The SSL expects to continually distribute white papers and notes aimed at both developers and users, and make documents available through the web.

A set of CD-ROM disks is being produced by the SSL and CSAT with sample Earth and space science data for use in testing and evaluating software. Most of NASA's science disciplines will be represented, most common data object types and formats will be included, and all of the datasets will have documentation and software for display. The SSL is considering producing a CD-ROM sampler disk with software developed under the AISRP, other popular application programs, and libraries of routines to support access to popular data formats.

Dr. Davis invited comments and suggestions regarding the initial version of the on-line SSL information service, as well as comments and suggestions on a software sampler CD-ROM. The SSL needs information for the abstracts on the AISRP projects and software.

The discussions focussed on the problems associated with software developed for different platforms, and what software can be supported in-house. Some of these problems may be conquered through use of a multi-platform software sampler.

SESSION I - ARTIFICIAL INTELLIGENCE TECHNIQUES

Chair: Dr. Richard Keller, Ames Research Center

Holographic Neural Networks for Space and Commercial Applications

Dr. Hua-Kuang Liu

Jet Propulsion Laboratory

The holographic neural network provides a pattern recognition system that can greatly amplify the target signal while filtering out the background noise and clutter. Partial input of very dimly illuminated targets can be recognized, and there are a number of NASA and commercial applications for this technology.

Dr. Liu discussed the approach and technical basis for the project, and provided a demonstration of the pattern recognition. Private industries have already expressed interest in the new

technology and end product of the project. Funding is provided by both NASA and ARPA. The feasibility study has been completed this year, and the prototype will be ready in 1995. If funding continues, this technology will be in users' hands in 2000.

The SEIDAM (System of Experts for Intelligent Data Management) Project

Dr. David Goodenough
Department of Natural Resources Canada

Intelligent data management systems are required to support the increased complexity and quantity of remote sensing data that is generated. SEIDAM, a NASA project supported by the Government of Canada and the Province of British Columbia, is a response to this need. The system integrates remote sensing data from satellites and aircraft with geographic information systems and manages large archives of remotely sensed data for query based recognition of forest objects appropriate for environmental forest monitoring. SEIDAM learns from examples or cases, consisting of a query, the data, and the desired goal.

Dr. Goodenough provided examples of queries and how SEIDAM answers those queries. The SEIDAM structure is a hierarchy of expert systems. Machine learning will be used to expedite the creation of knowledge and new expert systems and to learn from analysis. In order to construct the SEIDAM system, the project needed to accommodate different data sources--satellites, aircraft sensors, and field measurements. Dr. Goodenough showed how learning reduced the number of new rules (learning performance) for supervised classification, and displayed slides of sensor data and visualization examples. SEIDAM will connect to LandData BC, the future provincial land information system.

The first release of the Meta Data Model is complete, and the Meta Data Management subsystem is being developed now. In 1993 and 1994, satellite and aircraft acquisitions will take place over three test sites, and SEIDAM products will be demonstrated for these selected test sites.

After the presentation, the meeting participants discussed the SEIDAM level of accuracy, meta data, and utilization of commercial products.

SIGMA: An Intelligent Model-building Assistant

Dr. Richard Keller
Ames Research Center (ARC)

Over the past three years, the project goals have shifted. Initially, the project was to develop an intelligent domain-specific tool to assist planetary scientists in building scientific models. The

current goal has expanded to include the development of an intelligent generic tool to assist scientists in a variety of domains. SIGMA is a model building shell which consists of GUI, a scientific knowledge base, and knowledge acquisition tools. Scientists supply domain-specific knowledge and modeling scenario(s), and the product is a customized scientific modeling tool.

Dr. Keller addressed some of the difficulties inherent in model-building, and how SIGMA provides new tools to produce the code and facilitate the model building. SIGMA has been used to reproduce two models reported in scientific literature--Voyager I data analysis and forestry data. In addition, there have been numerous publications on the tool and a symposia. More work needs to be done on knowledge acquisition and maintenance, and control constructs need to be incorporated. Also, SIGMA cannot yet handle coupled equations, and for large data sets, a compiler needs to be built. The project is currently in an alpha test by the collaborators, and participants have been identified for beta tests. The project is also looking at moving into a new modeling domain--life support and exobiology.

Technology transfer plans include: a graduate student seminar at the University of Montana; a modeling seminar at the upcoming Ecological Society of America conference; a Cassini modelers meeting; and commercialization opportunities through the NASA Technology Applications Team.

In response to a question regarding how SIGMA compares with STELLA, Dr. Keller indicated that the types of models that can be built with STELLA are more restricted because knowledge of particular domains cannot be built into that system.

Multivariate Statistical Analysis Software Technologies for Astrophysical Research Involving Large Databases

Dr. George Djorgovski
California Institute of Technology

The project consists of the development of STATPROG, a user friendly, science-driven package for multivariate statistical analysis of relatively small data sets, and the development of SKICAT, an AI based system for processing and analysis of about 3 Tb of digital image information from the Second Palomar Sky Survey, and other present and future large astronomical data sets. SKICAT is a collaborative project between JPL and California Institute of Technology.

SKICAT consists of catalog construction elements (AutoPlate and AutoCCD) and external catalogs (IRAS, ROSAT, etc.). Catalog management is currently in progress, and the scientific analysis elements are yet to be developed. Dr. Djorgovski discussed the science drivers that require a cataloging of galaxies and point sources. The results will serve to validate galaxy theories and also provide a quality control check on the Space Telescope Science Institute. The major parts of SKICAT have been generated and tested, and some general utilities have been

developed. The results of the initial scientific verification and testing will be of great interest to astronomers. Future efforts include the completion of the catalog manipulation module and exploration of the high level machine learning and machine discovery techniques. Long term, the project aims to develop an evolving catalog through expert systems and machine learning techniques.

The focus of the discussion period for Session I was technology transfer. One of the challenges is how to get information on developing technologies out to the science community. The University of Maryland investigators were successful in using AGU this year, for two half-day sessions. When scientists start seeing published results using these tools, then the tools will move into the community more extensively. In certain discipline areas, such as astronomy and EOS, the old techniques will not work with the new orders of magnitude of the data sets, and effective tools will be necessary. The tool developer needs to be part of the "instrument builder" tradition. It is especially important that the scientists in the disciplines include attribution of tools in their publications and proposals.

At the AGU, there was a lot of interest in the tools, and what worked very well was demonstration of the tools by scientists instead of vendors. Based upon a survey of AGU attendees, the format of both oral presentation and demonstration was good.

Several ideas were suggested regarding ways to get published in mainline journals, including collecting several superior scientific papers and getting a special issue published.

After the discussion session, demonstrations of the AISRP projects tools and software were conducted. The items demonstrated are listed in Appendix D.

Wednesday, August 4, 1993

SESSION II - SCIENTIFIC VISUALIZATION

Chair: Dr. Theo Pavlides

The Grid Analysis and Display System (GrADS): A Practical Tool for Earth Science Visualization

Dr. James Kinter

University of Maryland

The GrADS was intended to be an Earth science tool, but is not strictly limited to weather graphics. The most important part of the GrADS design is integration--it integrates the manipulation of expressions and functions, access to data, and display in the form of maps, charts, and animation. It is interactive, easy to learn and use, and produces hardcopy via vector graphics.

Since last year, GrADS has been expanded in terms of data sets. It now supports GRIB and DRS as well as GrADS internal, and can support packed binary, ASCII, and netCDF. Currently, there are hundreds of GrADS users at over thirty institutions. It has been very successful for a number of reasons. It is easy to use and learn and it allows scientists to access and manipulate their data without additional programming. The software that is distributed is very simple for a scientist to install, and is platform independent--it runs on Unix workstations and PC's. GrADS has been distributed free of charge, and the Center has been willing to work with user groups to add the functionality they request and eliminate user problems. GrADS is used in research, in education, in forecasting, and for public information.

Dr. Kinter gave demonstrations of GrADS using others' software for a variety of research and forecasting applications. After the demonstrations and a brief video, Dr. Kinter described the major new features of GrADS, and what has been accomplished over the past year. Scripting language added during the past year permits users to customize all aspects of GrADS, and allows interaction with display for interactive scientific investigation using one's own data.

A Distributed Analysis and Visualization System for Model and Observational Data

Dr. Robert Wilhelmson

National Center for Supercomputing Applications (MCSA)

The project is part of a larger effort called PATHFINDER at NCSA. It is designed to bring to the researcher a flexible, modular, collaborative and distributed environment for use in studying

fluid flows, and one which can be tailored for specific scientific research and weather forecasting needs.

Since last year, a lot more work has been done with AVS. AVS is used to convert a subset of PATHFINDER modules on Convex and SGI machines. It is highly interactive, supports distributed computing, is extensible, provides user interface tools, and has a rich collection of existing modules. Other accomplishments over the past year include: the incorporation of GEMVIS, a subset of PATHFINDER which is a set of capabilities targeted to the GEMPAK user community; an animation tool; meta data; improved access to HDF; modification of some of the visual capabilities; laser disk recording; and prototyping and Explorer alpha and beta testing.

In September, the beta release of NCSA PATHFINDER Explorer modules will be made to the NCSA anonymous server, and GEMVIS will be released to Unidata. In October, NCSA plans the beta release of the Inventor animation tool, and release of GEMVIS to COSMIC. Final release of the NCSA Pathfinder Explorer modules is planned for summer of 1994.

In response to a question regarding large data sets and the difficulties in handling them with SGI, Dr. Wilhelmson agreed that there is need for a user tool that will do better memory management, which is very important for 3-D animation.

Handling Intellectual Property at UCAR/NCAR

Dr. Bill Buzbee

National Center for Atmospheric Research

There is a clear intellectual property issue whenever U.S. government funded work has gone abroad, and is used to develop products that are then sold back to the U.S. In addition, there is also an issue over how to take technology that has been developed with government funds and make it available to the private sector in a way that protects private investment and commercial viability. The UCAR/NCAR has developed an approach that addresses this issue.

The University Consortium for Atmospheric Research (UCAR) has a committee on intellectual property, and it works with the UCAR Foundation, a parallel organization that licenses technology that has commercial value. Dr. Buzbee described the intellectual property process for software. The developer discloses the intellectual property, and the UCAR Intellectual Property Committee assesses the market potential. If the committee determines that no market exists, the ownership is assigned to the developer or the public domain. If a market does exist, UCAR proceeds to copyright or patent the property and the National Science Foundation (NSF) obtains a nonexclusive and royalty-free license for government purposes. A dollar award is given to the employee(s), and the UCAR Foundation seeks a licensee. The license may be exclusive

or non-exclusive, depending on circumstances. A user license includes maintenance and user assistance via e-mail and phone, and the product is provided to universities at below cost. UCAR has found this to be an effective approach for handling technology transfer, and it provides protection to the value-added remarketer and for U.S. intellectual property.

NCAR Graphics--What's New in 3.2

Dr. Bob Lackman

National Center for Atmospheric Research

Dr. Lackman discussed the latest features of NCAR Graphics. Functionality has been improved through C language binding, raster contouring, high quality fonts, new documentation, utility upgrades, an interactive display tool (idt), and CGM translators. NCAR has also added some unsupported directories (Explorer modules and high level utilities). Version 4.0 will be out in late 1994 and will include the high level utilities, the NCAR command line (NCL) interpreter, NCL scripting language with loops and conditionals, a prototype GUI, and a user guide with tutorial and reference manual.

There were a number of questions related to the intellectual property portion of the NCAR presentation. The criteria for exclusive or non-exclusive licenses is primarily hardware driven. UCAR has occasionally gone outside for opinions, but typically the decision on whether property has market value is made by the Committee. If the product is not exclusively NCAR's, then legal agreements must be made with other parties before going through the licensing process. It is NCAR's policy that it will not include software unless these agreements are obtained. NCAR's ratio of success for commercially viable transfer is better than 50%.

McIDAS-eXplorer: A Version of McIDAS-X for Planetary Applications

Dr. S. S. Limaye

University of Wisconsin

The objective of the program was to adapt the McIDAS environment for analysis of planetary data for use on small computers/workstations so that it is useful for research, operations, and education. One of the most demanding challenges at this point in time is a support mechanism for users.

Currently, the project has a basic working core system on hand. No major problems have been encountered in the software development, with the exception of dynamic Magellan de-calibration. Most difficulties have been presented by the inability of different versions of Unix to read the PDS CD-ROM volumes. The incorporation of the NAIF/SPICE library in pre-existing software has progressed slower than anticipated, but no problems have been encountered.

The project plans to exhibit McIDAS-eXplorer at the annual meeting of the AAS/Division of Planetary Sciences in October. Dr. Limaye's group will experiment with some of the Mars Observer data when it is available for a trial run of use in a routine operational environment.

Dr. Limaye gave a demonstration of the program from a workstation. He demonstrated multi-frame display, query of data sets, animation, etc.

Experimenter's Laboratory for Visualized Interactive Science (ELVIS)

Dr. Elaine Hansen
University of Colorado

ELVIS is the integration of a number of technologies, tools, and theories to display and manipulate data. A core component of the system is PolyPaint, which enables 3-D rendering. The project has applied advanced human factors techniques and theories, including user interviews, cognitive walk-throughs, controlled user tests, and prototype evaluations. PolyPaint has been upgraded to handle new 2-D objects, and TAE+ has been used and augmented. ELVIS has extensible design which will support a scalable set of functions for shaded surface or volumetric renderings in real or indexed color.

Dr. Hansen demonstrated some of the features of ELVIS, including interactive user interface, 2-D and 3-D renderings of some typical space science data sets, the direct manipulation 3-D view and lighting editor, the intuitive color editor, and the PolyPaint rendering package.

Future plans include: integration with DataHub; integration of the spreadsheet engine; interface to graphics hardware via Inventor; integration of visualization of 2-D graphics in 3-D space; an alpha release in September; user interviews and user evaluations with their own data sets; and beta release at the end of the project.

Technology transfer has already been occurring in the development phase and will continue through the third year of the program. Dr. Hansen discussed a number of ways in which technologies have been or are being transferred--through transfer of advanced software tools and theories, transfer of intellectual resources, technology leveraging, evaluation testing, and technology insertion.

The general discussion for this Session focussed on hardware portability. The hardware base of the development community is not the same base as the science/university community. Is the next generation user community going to have SGI's, or are they going to stay with DOS/Windows systems? Some participants thought that tools and products should be developed for the next generation of systems, while others thought that tools should be written for what the

community has or will have. There was also some discussion on the issue of efficient use of resources. There are two modes of operation for most users: use of a high powered tool at a central location at the university, plus a variety of different types of hardware at the scientists' desks. Products need to address both modes, or have software that is scalable. As networks improve, some of these issues will be resolved, and data compression techniques may also help address the issue. Most participants felt that visualization packages should not be dependent upon a particular platform, but networks should be used to access resources wherever they exist. In general, users need tools that are modular, easy to use, and extensible.

SAVS: A Space and Atmospheric and Visualization System

Dr. E. Szuszczewicz

Science Applications International Corporation (SAIC)

The three major contributors to SAVS are the Laboratory for Atmospheric and Space Science at SAIC, Advanced Visual Systems (AVS), and the University of Maryland. SAVS is focussed on the multi-disciplinary databases designed to understand the cause-effect relationships in the solar-terrestrial system. It is composed of: widely accepted, commercially available visualization software with a heavily leveraged international users' module library; advanced distributed database techniques; and mathematical, analytical, and image processing tools.

In addition to the solar-terrestrial applications, SAVS is extensible to the lower atmospheric, Earth, and ocean sciences. SAVS is an integrated system that is easy to use, extensible, portable, financially accessible, and uses an end-to-end approach. Dr. Szuszczewicz discussed the seven major tasks in the project, the accomplishments over the past year, and future projected activities.

The SAVS demonstration focussed on the new, easier to use interactive front end. SAVS demonstrated an application using orbits, model, data, visualization, special algorithms, a 3-D to 1-D interpolation model, and remote data access. A special model has been developed to deal with the 3-D to 1-D interpolation. Dr. Szuszczewicz also demonstrated the ability to do remote data access realtime through NSSDC to data files in the SAIC laboratory. Activities in the third year will lead to full integration and tuning of all customized modules, and delivery of software and documentation, including the users' manual.

LinkWinds - The Linked Windows Interactive Data System

Dr. Amy Walton

Jet Propulsion Laboratory

LinkWinds provides a suite of tools to interactively visualize, explore, and analyze large, multivariate and multidisciplinary data sets. It will support the rapid prototyping and execution of data analysis and visualization, and allow maximum data and tools accessibility with a minimum of training.

During the past year, LinkWinds version 1.4 was released to a significantly expanded number of scientific groups at a variety of institutions. New applications were added and existing applications expanded. In addition, the project implemented a realtime data interface to LinkWinds and used it to support the University of Iowa's Plasma Wave Spectrometer aboard the Galileo spacecraft.

To facilitate technology transfer, the project has made ingestion of data as easy as possible. Direct LinkWinds interfaces have been provided for key datafile formats (HDF, CDF, and RAW), and arrangements have been made to provide DataHub. All new users will be provided with an interactive realtime tutorial using MUSE.

Dr. Walton conducted a brief demonstration of LinkWinds. Currently, there are ten sites actively using LinkWinds in research, and feedback from scientists has been very positive. In response to a questions, Dr. Walton indicated that LinkWinds uses straightforward grid data, and although it handles 3-D data sets, it does not do data rendering.

DataHub

Dr. Tom Handley

Jet Propulsion Laboratory

DataHub is a value-added, knowledge-based server between the data suppliers and the data consumers. DataHub understands the different scientific data models, and addresses a variety of needs. Dr. Handley described the functional architecture of DataHub, as well as the data model.

Initially, the project attempted an artificial neural network, but is now investigating machine learning techniques for feature detection and learning. Although DataHub was developed with an emphasis on physical oceanography, it is now processing and extracting information from multi-spectral data. DataHub is being used (and delivered) with other applications, such as LinkWinds and PolyPaint.

Dr. Handley walked through a demonstration of the DataHub input, process, and output modes. In the near term, the project will be working on a journal and transaction manager, a more general data model for user-defined data and conversion, and an expert system for dataset type checking. Long term work will include a context-based and content-based data search capability, quality control and "corporate" memory of the conversion of input data points to output data points, and a more general search path and browsing mechanism. In response to a question, Dr. Handley indicated that the project is now looking at how to handle non-gridded data. The interface could be broadened to include connection to a visualization tool, but it would need a data model big enough to handle complicated, non-gridded data sets. Also, the project is trying to make DataHub independent of any tool or database.

Data Visualization and Sensor Fusion

Dr. Vance McCollough

Hughes Aircraft Company

This project, a collaboration between the University of Chicago and Hughes, is aimed at the application of medical imaging and visualization techniques to remote sensing. The DMSP data set was selected for study. Image processing and visualization algorithms were developed and tested on CAT and PET medical data. This included image matching using landmarks, contouring and thresholding, image linking, 3-D image processing/visualization and superposition of 2-D images, and image-indexed color table generation.

Dr. McCollough described the characteristics of the DMSP data set, and how calibration, navigation, projection, and gridding was accomplished. Medical imaging software was not as well suited to projection and gridding problems as had been anticipated. Dr. McCollough showed examples of thresholding sensitivity, image-based color table indexing, and three-dimensional visualization of IR images. The work done at Hughes showed the results of how the principal components transformation of microwave channels can extract useful data, and the principal components can be used to drive hue, saturation, and intensity to produce useful visualizations of multi-channel data.

For the remainder of 1993 and 1994, the project plans to: apply visualization algorithms to remote sensing data; acquire additional data for evaluation; evaluate visualization results; investigate applications of principal component techniques; and do additional work on 3-D visualization of IR images.

In the general discussion on Session II, the primary topic was meta data and data description. One observation made was that some tools are very useful, but are based on uniformly gridded data. Some space science is not image-based, and the issue is whether or not these tools can be

adapted for those purposes. There was general agreement that data description needs to be well documented, precise, and concise. If data can be described in a complete and concise manner, the need for data standards is not as critical. What is needed is a standard framework in which to provide data information--a "format of formats." In addition, data structure needs to be considered separately from data format. There was discussion on the advantages and disadvantages of general use of tools vs. development of tools for specific purposes. Developing the tools, using them intelligently, and understanding the data are all separate issues.

Thursday, August 5, 1993

SESSION III - DATA COMPRESSION/ARCHIVING/ACCESS/ANALYSIS

Chair: Dr. James Storer

Data Compression

Dr. James Storer
Brandeis University

Data compression research at Brandeis includes lossless compression, image compression, video compression, and error resilient compression. This presentation focussed on on-line image compression with adaptive vector quantization (VQ) with variable sized vectors. VQ is a lossy technique which has not been used extensively in the past due to its static characteristics. However, new work on a version of VQ overcomes this disadvantage. The on-line adaptive VQ is characterized by an "evolving" dictionary.

Dr. Storer described how the generic encoding algorithm works. Wave has proved to be the most efficient growing strategy, and all of the results shown in the presentation used the wave strategy. Dr. Storer also described how the growing points were selected, and how the best rectangle is chosen from the dictionary. The dictionary was modified by growing larger rectangles from smaller ones used in the past. The test sequence included a variety of images. Three experimental runs were done for each file, and compression levels were compared with JPEG. JPEG is tuned for magazine photos, and JPEG compression was superior on those images. The adaptive VQ compression was tuned for more unusual images.

A video compression approach is being developed using a superblock displacement estimation technique. The superblock technique gives a factor of three to four improvement over fixed size blocks. For science, a lot of images can be treated as video because the sequences look like video and displacement estimation techniques are very effective. The goal is realtime high

fidelity video and image sequence compression of over 1000 to one. The project expects to report more on this area next year.

SAMS: A Spatial Analysis and Modeling System for Environmental Monitoring

Dr. Fran Stetina

Goddard Space Flight Center

The objective of this project was to develop an end-to-end system to support environmental monitoring and management. It was a low-end pilot study for EOS direct readout. SAMS has incorporated realtime data reception, multi-discipline analysis and modeling, and expert systems. Dr. Stetina emphasized that this is not "cutting-edge" technology, but an attempt to adapt and integrate some new tools into existing technology that is reliable and maintainable. Most of the SAMS activities have been involved in integrating elements that already exist. The front end of the system can take any satellite data, realtime, and is automatic and transparent to the user. SAMS is multi-discipline--it can handle atmospheric, land, and ocean processes. It has modules that will analyze data to support marine resources, water resource management, disaster management, and agriculture, as well as environmental monitoring. SAMS has a built-in DMS, performs data fusion, and is involved in distributing products through Internet, Peacesat, and Earth Alert, as well as hard copy and storage.

SAMS, or a subsystem of it, has been utilized in ten projects or field operations centers around the world. SAMS started in 1988 with a direct readout, and will move into EOS direct readout in 1995. Dr. Stetina described the accomplishments at the agricultural and environmental sites around the world that utilize SAMS. The software is available for users that have the hardware. In response to a question, Dr. Stetina indicated that the project wants to do more work with data compression, and this has been looked at for some of the data.

Land-Surface Testbed for EOS

Mr. Tim Kelly

University of Colorado

Mr. Kelly described the historical background of Sanddunes, a land-surface image system for Earth resources. Current work in process includes the development of a stand-alone data archive system, the addition of DOMSAT data, and the addition of GrADS overlays. The total number of registered users (based upon login records) has risen to over 2500 as of June 1993, and the project is shipping about 250 images per month.

The new element of the system, called Navigate, enables Sanddunes to send a navigate command to an IBM 6000 cluster (which navigates the images) and receive images into the user file. The system is designed to be totally self sufficient. Plans are to expand the present system to include global AVHRR data via the DOMSAT antenna, to assess the capability of the new independent storage system to handle the greater volumes of DOMSAT and other satellite data, and to explore the capacity of the new network capabilities to handle the transfer of all stored and newly collected data. This testbed will transition to the Version 0 EOSDIS before December 1994, and will be located at either JPL or GSFC, or both. Until then, the Sanddunes will reside at LASP. Although the current system is free, a potential user must be on Internet.

Development of a Tool-Set for Simultaneous, Multi-Site Observations of Astronomical Objects

Dr. S. Chakrabarti
Boston University

The primary purpose of this project was for education--to make the system available to all astronomers, both amateur and professional. There will be a large number of users connecting through Internet to robotic astronomical telescopes. Dr. Chakrabarti described the present telescope control configuration. Upon receiving a telescope request, the telescope management program uses a telescope database to determine if a suitable telescope is available and/or if the object requested is observable. If no suitable telescope is available or the observation is not possible, an error message is relayed back to the user. If a telescope is found and observation is possible, a telescope interface is established and the user starts observation.

Dr. Chakrabarti described several elements of the system which are used to link the user to one telescope at a time, allow simultaneous requests to more than one telescope from a single user, and allow monitoring of an observation in progress. There are three types of users of this system--resident observers, remote observers, and remote watchers (who have no interactive capability). The project is continuing to work networking issues (robustness, fault tolerance, security, and data compression) and software issues (portability, support, and public domain). Currently, all of the hardware and networking tools are in place, and some user interface has been written. Future activities include: improvements to the user interface; realtime tests; the addition of imaging, spectroscopy, and interferometry capabilities; the addition of other platforms; and the addition of analysis and artificial intelligence tools.

The general discussion session focussed on the topic of archiving data. The Workshop participants discussed the pros and cons of letting market needs drive the nature and extent of data archive. For the Global Change Research Program, the government needs to have access to historical environmental data, and a case was made that the government should be involved in the archive of this type of data. However, it was recognized that currently there is not an

infinite resource for keeping data, and that sensors have the capability of generating more data than can possibly be stored. One idea presented was to archive the raw data, but not all of the data products. In the future, with new state-of-the-art media, storage may not be as much of an issue as access.

Geographic Information System (GIS) for Fusion and Analysis of High Resolution Remote Sensing and Ground Truth Data

Dr. Anthony Freeman

Jet Propulsion Laboratory

The overall aim of the project was to expand the user community for radar images by simplifying data display, analysis, and interpretation. The project started out by adapting a VICAR/IBIS GIS including models and different data sets, especially radar images. It achieved a working version of the system with a peer user interface. Other achievements included the development of MacSigma 0, work on supervised classification of radar images, work on validation models, the development of a three-component scattering model, and the development of a vegetation map classification (MAPVEG) expert system.

Dr. Freeman described the three-component scattering model and several scattering mechanisms. The project developed a technique of estimating the contribution of each of the three mechanisms and tried to develop an overall classifier. They developed some classification rules, as part of the expert system, and came up with a classifier for the three images trained on (Netherlands farmland, a rain forest, and Black Forest/farmland). The classifier was then run with other images, and did a very good representation. Unsupervised classification can be applied to any calibrated three-frequency AIRSAR data, but it is not designed for very low incidence angles.

Dr. Freeman discussed the software that has been developed in this program. MacSigma 0, released through COSMIC, is for display and analysis of radar images and has export capabilities designed into the software. MAPVEG, which contains the classifier, takes AIRSAR data and produces a vegetation map. RAVEN, currently under development, performs display and analysis of radar images on UNIX. Image registration software is planned. MAC software and documentation is to be incorporated into the SIR-C education CD.

Envision: A System for Management and Display of Large Data Sets

Dr. Kenneth Bowman
Texas A&M University

Envision integrates data management, manipulation, analysis, and display functions into a single interactive environment. It uses standard portable data storage and management tools to provide access to multi-GB data sets and provides a simple, intuitive, collaborative, and portable user interface. Envision will enhance the capabilities of existing interactive visualization software developed at the NCSA.

The Envision Data Manager consists of a data server and maintains the project file and data storage. The user interface provides a table-like display of meta data. The initial release uses NCSA XImage, NCSA Collage, and IDL. The project is planning to complete some functions in the Data Manager, but is now focussing on making it easier for people to connect their own tools and software.

Dr. Bowman gave a brief demonstration of Envision, which is available through anonymous ftp.

The meeting adjourned to the demonstration session.

Friday, August 6, 1993

SESSION IV - RESEARCH AND TECHNOLOGY

Chair: Dr. Amy Walton

Research and Technology Activities at JPL

Dr. Amy Walton
Jet Propulsion Laboratory

Dr. Walton discussed the current research and technology activities at JPL. The scope of the activities include: the development and exercise of tools that make use of visual perception to integrate and display multiple diverse data sets, explore and validate data sets, track and measure features and dynamics, provide perspective simulations, and generate scientist controlled animations; the automation and acceleration of analysis of large complex data sets; the development of interactive capabilities; the extension of capabilities to a heterogeneous computing environment; the incorporation of techniques that make the environment self-training and extensible; and the migration of developed tools into the science environment.

JPL is currently working on improvements in data access, such as data storage options, ancillary data, and standards/data interoperability. JPL also is working on integration and visualization testbeds and techniques that specifically support data analysis. Some of these projects are: imaging methods for multi-dimensional data visualization, graphical methods for science data analysis, scientific tools for JPL image archives, and an integrated science analysis testbed. Dr. Walton discussed each of these activities.

Future directions are towards "desktop science", which will be a move from timesharing/LAN to WAN and will involve the use of smaller local machines to access varied computing resources with the ability to use varied data formats, multi-vendor hardware, and software. The ability to rapidly evaluate large data volumes, complex data sets, and combinations of data sets has become a major need in the scientific research process. Dr. Walton emphasized the importance of "people transfer", or working together, as a key aspect of technology transfer.

Navigation Ancillary Information Facility: An Overview of SPICE

Dr. Amy Walton

JPL

There are two kinds of space science data--science instrument data and ancillary engineering data. SPICE (Spacecraft Planet Instrument C-matrix Events) addresses ancillary engineering data, which can be from the spacecraft, mission control center, and scientists. Ancillary data are data that tell when an instrument was taking data, where the spacecraft was located, how the spacecraft and its instruments were oriented, and what was the size, shape, and orientation of the targets being observed. It can also tell how the instrument was acquiring the data and what else was happening on the spacecraft or in the ground data system.

The principal SPICE system components are data files and software, plus standards, documentation, user support, and system maintenance. It was originally intended for space science data analysis, but is now also being used for science observation planning and mission evaluation from a science perspective. SPICE is currently being used on Voyager, Magellan, and Galileo. It is being planned for Hubble Space Telescope and other flight project and terrestrial programs. Possible future applications include MESUR, the Discovery Program, AXAF, EOS, and INTERBOL. In addition, SPICE can be used as a component of a space science education program at the university level.

Using the NAIF SPICE Kernel Concepts and the NAIF Toolkit Software for Geometry Parameter Generation and Observation Visualization

Dr. Karen Simmons

LASP, University of Colorado

Dr. Simmons provided some historical background on why a program like SPICE was needed and how SPICE is being used. Better navigation data was needed because the Supplemental Experiment Data Records (SEDRs) were static products and had conceptual flaws with respect to spacecraft navigation, pointing, and science instrument design improvements. A working group came up with a better way of doing this: SPICE Kernels--independent pieces of knowledge that would contain 99.9% of the information that is wanted on what the data is and how the data is being collected. Dr. Simmons discussed the advantages of the SPICE Kernels. Kernels created the need for standardized parameter definition and generation. The SPICE Toolkit helps uniform understanding and use of geometry parameters.

The Geometry and Graphics Software (GGS) provides the expertise to use both the Kernels and the Toolkit. GGS is a tool which allows the scientist to understand the complex geometric environment in which a data set was obtained. GGS provides geometry parameters via active display, hardcopy, or footprint files, and provides science observation visualization via after-the-fact animation of the commanded sequence, opportunity investigation, and planning and design. GGS maintains and documents expertise both actively (mostly through Kernels) and passively.

Two types of Kernels represent all the knowledge needed about how data was collected. The SPICE Kernels are either I-Kernel (instrument related) or E-Kernel (event, experimenter, or expert system related). SPICE is an exceptionally viable system. However, the Kernel generation history documentation needs improvement (guidelines are needed), and the E-Kernels, although they represent a vital link, are still in development.

The project has developed a number of versions of GGS, and has learned enough about how it works to make it a multi-mission tool. However, at present there is no way of handling differences among platforms. IDL was chosen for the scientific interaction with data and geometry because copies need to be propagated across different versions. However, as new IDL versions are released, the software must be kept updated. Future work can include: more design-side tools; C-Kernel Smithing with non-imaging data; tour analysis tools; and investigation of other uses, such as EOS, a PDS archive tool, and marriage with a sequence planning tool like OASIS.

A captive guest account has been established to provide a demo of the GGS software package. Details are included in the presentation material in Appendix E.

The EOS/Pathfinder Interuse Experiment

Dr. Mike Botts

University of Alabama in Huntsville

Interuse signifies the ease with which data sets from various sensors and disciplines can be brought together, coregistered in space and time, correlated, analyzed, and visualized together within scientific tools. The present EOS direction of relying heavily on previously gridded and projected Level 3 data is inadequate for several reasons. Dr. Botts described the experiment plan. The new approach, Interuse, focusses on making Level 2 data more useful. However, several issues must be resolved before an easily used Level 2 data set can be produced and distributed: generic navigation for EOS data; generic gridding and projection routines; incorporation of navigation information into HDF; data subsetting and compression; and the availability and compatibility of analysis and visualization tools. A number of key pieces could help with this experiment: SPICE, SPICEb and OoSPICE, PLATO, and extensible visualization tools such as LinkWinds, IDL/PV-Wave, AVS, Explorer, and Khoros. The project is looking at a modification of some of these tools that will allow navigation within the tool itself.

Two Interuse Teams have been established. The Interuse Tiger Team will be developing some of the elements and focussing the activity. The Interuse Core Working Group, which includes scientists and data producers, will provide feedback, checkout, test, etc. Dr. Botts discussed the schedule for the experiment. During Phase I, the project will work through the issues. Phase II will consist of generation of data sets and distribution. Phase III will include assessment and refinement.

Overview of Ames Research Center Advanced Network Applications

Dr. John Yin

Ames Research Center

Dr. Yin described the functions of several elements of the ARC organization that are involved in advanced network applications. The Advanced Network Applications (ANA) Section at ARC provides network based solutions for NASA and other federal agencies through the development, implementation, and operations coordination and support of extensible high technology network services and applications. The Network Applications Information Center (NAIC), part of the ANA section, provides NASA-wide operations support of advanced network services for the NASA science and research community. The Network Services Development Group develops, implements, and transitions to the NAIC state-of-the-art standards-based network applications and services.

Dr. Yin discussed some of the ANA development activities in detail. The X.400/SMTP electronic mail gateways provide an agency wide electronic messaging infrastructure which

allows for completely transparent document exchanges between all host and LAN based mailers. The distributed X.500 directory service allows for the management and distribution of an electronic locator service common to all NASA sites. The electronic signatures and certificate management is a joint activity with the NASA Headquarters Office of the Comptroller, NIST, and USPS to develop and demonstrate the generation, certification, management, and use of electronic signatures for authenticating electronic documents. Privacy enhanced mail provides a standards based end-to-end encryption of electronic messages. Packet video development and deployment is a project sponsored by the NASA Headquarters Office of Aeronautics HPCC program for providing interoperable color packet video to UNIX workstations, PCs, and Macs by the end of FY 1994, with initial demonstration on Sun workstations by the end of FY 1993. Wireless network development will demonstrate high speed wireless connectivity to portable computers with special emphasis on support for advanced applications like packet video.

The participants discussed network accessibility, particularly Internet, which is still not widely used in the administrative and business sectors. Other issues discussed were funding for connection to Internet for the science community, and a "dial-up" capability for NASA science Internet. Mr. Mucklow emphasized that NASA encourages the use of networks for science collaboration, as reflected in the latest NASA Research Announcement (NRA). The NRA was issued electronically, and the evaluation process has been set up electronically. When NASA has electronic signature capability, proposals may be submitted electronically as well. Mr. Mucklow noted that all electronic proposal documentation should be in Postscript or Acrobat rather than ASCII format.

Summary and Action Items

Mr. Mucklow invited Dr. Jim Dodge from the Mission To Planet Earth (MTPE) Office to share some of his thoughts on the AISRP and the Workshop. Dr. Dodge indicated that there were a lot of good elements demonstrated at the Workshop that he will be taking back to share with the Global Change Data Analysis Program. He particularly praised the tools for data handling/merging and tools for visualization and data fusion. He suggested that an article be written summarizing the tools that have been developed or are being developed. Accessing and analyzing extremely large amounts of data will be needed in EOS, and this aspect did not appear to be addressed fully in the Workshop presentations and demonstrations. Other challenges which need to be addressed are a comprehensive approach to error tracking and accuracy (e.g., an error tracking system on a pixel basis), and data compression, which will continue to be an issue with researchers. Dr. Simmons suggested that the process that the Galileo scientists had to go through in making decisions regarding compression (due to the antenna problem) might be of value to the EOS scientists.

One of the advantages of the AISRP is that it is multi-disciplinary, and scientists have an opportunity to learn from each other and exchange tools and software. Tools will often reveal problems in data and models, and programs need to take care that the data sets and models that are generated have high quality control. The desktop computer to supercomputer via networks is the new paradigm, and requirements are increasing while resources are decreasing.

The group discussed the challenges in getting new tools into the hands of more scientists. There was consensus that this needs to be pushed more through the discipline technical societies, and this action was given to the scientific participants. A key element is to select partner users that are well respected and produce first-class science. If the groups using the tools are well respected in the science community and produce notable science results, then others in that community will be highly motivated to use the tools.

One of the issues is support for users when a tool becomes successful. Mr. Bredekamp indicated that in the current environment of shrinking resources, developers need to find clever ways to minimize the need for support, and design tools that are very easy to use. NASA and the science community cannot continue to do software maintenance as has been traditionally done. One positive aspect that will help with this challenge is the continued collaboration and union of the computer science community with Earth and space scientists, which has been emphasized in the AISRP.

Other specific needs include more work in the algorithm area, and tools that use AI techniques to search data in extremely large data sets. Before adjourning the Workshop, Mr. Mucklow noted that the ISB has the action item to send out the Integrated Technology Plan to the Workshop attendees, and the AISRP investigators have the action to get started earlier this year with the AGU on a special session.

APPENDIX A

AISRP Agenda

**APPLIED INFORMATION SYSTEMS RESEARCH PROGRAM (AISRP)
WORKSHOP III
AUGUST 3-6, 1993
Laboratory for Atmospheric and Space Physics (LASP)
1234 Innovation Drive
University of Colorado, Boulder, CO
303/472-6867 fax 303/472-6444**

Meeting Chair: Mr. Glenn H. Mucklow/NASA Office of Space Science

Tuesday, August 3

- 7:30 Coffee and Registration
- 8:30 Opening Remarks
Mr. Joseph Bredekamp/NASA Office of Space Science
- 9:00 Software Support Laboratory and data Sampler CD-ROM
Dr. Randy Davis, University of Colorado

SESSION I - ARTIFICIAL INTELLIGENCE TECHNIQUES

Chair: Richard Keller

- 9:30 Announcements
- 9:45 Multi-Layer Holographic Bifurcative Neural Network Systems for Real-Time Adaptive EOS Data Analysis
Dr. Hua-Kuang Liu, JPL
- 10:15 Break
- 10:30 System of Experts for Intelligent Data Management (SEIDAM)
Dr. David Goodenough, Energy Mines & Resources
- 11:00 Construction of an Advanced Software Tool for Planetary Atmospheric Modeling
Dr. Richard Keller, ARC
- 11:30 Multivariate Statistical Analysis SW Technologies for Astrophysics Research Involving Large Data Bases
Dr. George Djorgovski, California Institute of Technology
- 12:00 Lunch
- 1:30 SESSION I DISCUSSIONS
- 2:30 DEMONSTRATIONS
- 3:00 Break
- 5:00 ADJOURN to Reception
- 6:30 Reception at NCAR

Wednesday, August 4

8:00 Coffee

SESSION II- SCIENTIFIC VISUALIZATION Chair: Theo Pavlidis

8:30 Announcements

8:45 The Grid Analysis and Display System (GRADS): A Practical Tool for Earth Science Visualization
Dr. James Kinter, University of Maryland

9:15 A Distributed Analysis and Visualization System for Model and Observational Data
Dr. Robert Wilhelmson, NCSA

9:45 An Interactive Interface for NCAR Graphics
Dr. Robert Lackman, NCAR

10:15 BREAK

10:30 Topography from Shading and Stereo
Dr. Berthold Horn, MIT

11:00 Planetary Data Analysis and Display System: A Version of PC-McIDAS
Dr. Sanjay Limaye, University of Wisconsin

11:30 SESSION II DISCUSSIONS A

12:00 LUNCH

1:30 Experimenter's Laboratory for Visualized Interactive Science
Dr. Elaine Hansen, University of Colorado

2:00 SAVS: A Space Analysis and Visualization System
Dr. Edward Szuszczewicz, SAIC

2:30 LinkWinds: A Distributed System for Visualizing and Analyzing Multivariate and Multidisciplinary Data
Dr. Allan Jacobson, JPL

3:00 BREAK

3:15 DataHub: Knowledge-based Assistance for Science Visualization and Analysis Using Large Distributed Databases
Mr. Thomas Handley, JPL

3:45 Advanced Data Visualization and Sensor Fusion
Mr. Vance McCullough, Hughes Aircraft Co.

4:15 SESSION II DISCUSSIONS B

5:00 ADJOURN

Thursday, August 5

8:00 Coffee

**SESSION III - DATA MANAGEMENT AND ARCHIVING:
DATA COMPRESSION/ARCHIVING/ACCESS/ANALYSIS**

Chair: Dr. James Storer

8:30 Announcements

8:45 High Performance Compression of Science Data
Dr. James Storer, Brandeis University

9:15 VIEWCACHE: An Incremental Pointer based Access Method for
Autonomous Interoperable Databases
Dr. Nicholas Roussopoulos, University of Maryland

9:45 A Spatial Analysis and Modeling System for Environment Management
Mr. Fran Stetina, GSFC

10:15 BREAK

10:30 A Land-Surface Testbed for EOSDIS
Dr. William Emery, University of Colorado

11:00 Development of a Tool Set for Simultaneous Multi-Site Observations of
Astronomical Objects
Dr. Supriya Chakrabarti, Boston University

11:30 SESSION III DISCUSSIONS A

12:00 LUNCH

1:30 Geographic Information System for Fusion and Analysis of High
Resolution Remote Sensing and Ground Truth Data
Dr. Anthony Freeman, JPL

2:00 An Interactive Environment for the Analysis of Large Earth Observation
and Model Data Sets
Dr. Kenneth Bowman, Texas A&M

2:30 SESSION III DISCUSSIONS B

3:00 DEMONSTRATIONS

5:00 ADJOURN

Friday Morning, August 6

8:00 Coffee

SESSION IV: RESEARCH AND TECHNOLOGY

Chair: Dr. Amy Walton, JPL

8:30 Announcements

8:45 Overview of ISB Research and Technology Activities
Dr. Amy Walton, JPL

9:15 Navigation Ancillary Information Facility
Dr. Chuck Acton, JPL

9:45 Using the NAIF SPICE Kernel Concepts and the NAIF Toolkit Software
for Geometry Parameter Generation
Dr. Karen Simmons, University of Colorado

10:15 Break

10:30 Advanced Network Applications
Mr. John Yin, ARC

11:00 Splinter Group Reports

11:30 Summary and Action Items

12:00 END OF WORKSHOP III

APPENDIX B

Participants & Attendees

ATTENDEES AISRP WORKSHOP III AUGUST 3-6, 1993

Acton
 Jet Propulsion Laboratory
 Pasadena, CA 91109

Chuck

Alpert
 National Center for Atmospheric Research
 P.O. Box 3000
 Boulder, CO 80307

Ethan

303/497-1832

FAX: 303/497-1137

Blanchard
 Science Applications International Corporation
 1710 Goodridge Drive
 MS T-2-3-1
 McLean, VA 22102
 703/556-7108
 blanchard@mclapo.saic.com

Paul

FAX: 703/821-1134

Botts
 University of Alabama, Huntsville
 Earth System Science Lab
 RI A-11
 Huntsville, AL 35899
 205/895-6257
 botts@cyclone.msfc.nasa.gov

Mike

FAX: 205/895-6970

Bowman
 Associate Research Scientist
 Climate System Research Program
 Texas A & M University
 College Station, TX 77843-3150
 409/862-4060
 bowman@csrp.tamu.edu

Kenneth P.

FAX: 409/862-4132

Bredekamp
 Chief, Information Systems Branch
 Code ST
 National Aeronautics & Space Administration
 Washington, DC 20546
 202/358-2473
 JBredekamp@nasamail.nasa.gov

Joseph H.

FAX: 202/358-4166

Bruehl
 Unidata Program Center
 UCAR
 PO Box 3000
 Boulder, CO 80307
 303/497-8641

Peggy

FAX: 303/497-8690

Buzbee
 Director
 National Center for Atmospheric Research
 P.O. Box 3000
 Boulder, CO 80307
 303/497-1206
 buzbee@ncar.ucar.edu

William

FAX: 303/497-1137

ATTENDEES AISRP WORKSHOP III AUGUST 3-6, 1993

Chakrabarti Supriya
 Center for Space Physics
 Boston University
 725 Commonwealth Avenue, Room 506
 Boston, MA 02215
 617/353-5990 FAX: 617/353-6463

Condrescu Michael
 Space Environment Laboratory
 NOAA
 325 Broadway
 Boulder, CO 80303

Cruickshank Cheryl M.
 Space Environment Laboratory
 NOAA R/E/SE
 325 Broadway
 Boulder, CO 80303
 303/497-3930 FAX: 303/497-7392
 ccruickshank@selvax.sel.blrdoc.gov

Davis Randall
 Laboratory for Atmospheric & Space Physics
 University of Colorado
 Campus Box 392
 Boulder, CO 80309
 303/492-6867 FAX: 303/492-6444
 AQUILA::DAVIS

Djorgovski George
 Department of Astronomy
 M.S. 105-24
 California Institute of Technology
 Pasadena, CA 91125
 818/395-4415 FAX: 818/568-9352
 George@deimos.caltech.edu

Dodge James C.
 Code YS
 National Aeronautics and Space Administration
 Washington, DC 20546

 202/358-0763 FAX: 202/358-3098
 JDodge

Doty Brian E.
 Department of Meteorology
 University of Maryland
 College Park, MD 20742

 301/405-5356 FAX: 301/314-9482
 doty@cola.umd.edu

Dueck Sandy
 MS 168-514
 Jet Propulsion Laboratory
 4800 Oak Grove Road
 Pasadena, CA 91109
 818/354-5073 FAX: 818/393-6962
 sdueck@jplpds.jpl.nasa.gov

ATTENDEES AISRP WORKSHOP III AUGUST 3-6, 1993

Eischeid
CIRES

J. F.

Emery William
Colorado Center for Astrodynamics Research
University of Colorado
CB 341
Boulder, CO 80309
303/492-8591 FAX: 303/492-2825
bemery@nasamail.nasa.gov

Evans Dave
Space Environment Laboratory
NOAA
325 Broadway
Boulder, CO 80303

Freeman Anthony
Jet Propulsion Laboratory
M.S. 300-235
4800 Oak Grove Drive
Pasadena, CA 91109
818/354-1887 FAX: 818/393-6943
freeman@jplmrs.jpl.nasa.gov

Fry Patrick
Space Science and Engineering Center
University of Wisconsin-Madison
1225 W. Dayton Street
Madison, WI 53706
608/263-9694
fry@miranda.ssec.wisc.edu

Fung Ko B.
Canada Centre for Remote Sensing
Department of Natural Resources
588 Booth Street
Ottawa, Ontario KIA OY7 CANADA
613/947-1234 FAX: 613/947-1383
fung@cers.cmr.ca

Galucci Michael
Space Environment Laboratory
NOAA R/E/SE
325 Broadway
Boulder, CO 80303
303/497-7448 FAX: 303/497-3645
mikeg@sel.bldrdoc.gov

Goodenough David G.
Pacific Forestry Centre
Department of Natural Resources
506 West Burnside Road
Victoria, B.C. CANADA V8Z1M5
604/363-0776 FAX: 604/363-0775
dgoodenough@A1.pfc.forestry.ca

ATTENDEES AISRP WORKSHOP III AUGUST 3-6, 1993

Goodrich Charles C.
 Astronomy Department
 University of Maryland
 College Park, MD 20742

301/405-1516 FAX: 301/314-9966
 ccg@alv.umd.edu

Greer Sue
 Space Environment Laboratory
 NOAA
 325 Broadway
 Boulder, CO 80303

Hagedorn John
 Code 912
 Goddard Space Flight Center
 National Aeronautics and Space Administration
 Greenbelt, MD 20771
 301/286-7374 301/286-4661
 hagedorn@betsy.gsfc.nasa.gov

Handley, Jr. Thomas J.
 M.S. 168-414
 Jet Propulsion Laboratory
 4800 Oak Grove Drive
 Pasadena, CA 91009
 818/354-7009 818/393-6962
 THandley@spacemouse.jpl.nasa.gov

Hansen Elaine
 Colorado Space Grant College
 University of Colorado
 Box 520
 Boulder, CO 80309
 303/492-3141 FAX: 303/492-5456
 hansen@vega.colorado.edu

Heckman Gary
 Space Environment Laboratory
 NOAA R/E/SE
 325 Broadway
 Boulder, CO 80303
 303/497-5687 FAX: 303/497-7392
 gheckman@selblrdoc.gov

Hill John
 Houston Advanced Research Center
 4800 Research Forest Drive
 Houston, TX 77381

713/363-7999 FAX: 713/263-7931

Hillis Eric
 LASP Space Technology Building
 University of Colorado
 1234 Innovation Drive
 Boulder, CO 80303
 303/492-7289 FAX: 303/492-6444

ATTENDEES AISRP WORKSHOP III AUGUST 3-6, 1993

Hughes Peter
 Code 522.3
 Goddard Space Flight Center
 Greenbelt, MD 20771
 301/286-3120 FAX: 301/286-4627
 phughes@kong.gsfc.nasa.gov

Jacobson Allan
 Jet Propulsion Laboratory
 M.S. 183-501
 4800 Oak Grove Drive
 Pasadena, CA 91109
 818/354-0692 818/354-0966
 BudJ@apex.nasa.gov

Keller Richard
 AI Research Branch
 Ames Research Center
 M.S. 269-2
 Moffett Field, CA 94035-1000
 415/604-3388 FAX: 415/604-3594
 keller@ptolemy.arc.nasa.gov

Kelley Timothy
 Colorado Center of Astrodynamics
 CB 431
 Boulder, CO 80303
 303/497-1221 FAX: 303/497-2825
 kelley@sanddunes.scd.ucar.edu

Kinter James L
 COLA
 Institute of Global Environment
 4041 Powder Mill Road, Suite 302
 Calverton, MD 20705

Klemp Joseph
 National Center for Atmospheric Research
 PO Box 3000
 Boulder, CO 80307
 303/497-8902 FAX: 303/497-8181
 klemp@ncar.ucar.edu

Klemp Marjorie
 LASP Space Technology Building
 University of Colorado
 Box 590
 Boulder, CO 80309
 303/492-7289 FAX: 303/492-6444
 Margi@rana.colorado.edu

Krauss Robert
 Space Science and Engineering Center
 University of Wisconsin-Madison
 1225 W. Dayton Street
 Madison, WI 53706
 608/262-9573
 bobkr@ssecmail.ssec.wis.edu

ATTENDEES AISRP WORKSHOP III AUGUST 3-6, 1993

Kulkarni Ravi
Advanced Visualization Lab
University of Maryland
College Park, MD 20742

301/405-1559 FAX: 301/405-9966
ravi@avl.umd.edu

Kunches Joe
Space Environment Laboratory
NOAA R/E/SE
Boulder, CO 80303

Lackman Robert
National Center for Atmospheric Research
1850 Table Mesa Drive
Boulder, CO 80307-3000

303/497-1224 FAX: 303/497-1137
RLL@ncar.ucar.edu

Limaye Sanjay S.
Space Science and Engineering Center
University of Wisconsin
1225 W. Dayton Street
Madison, WI 53706
608/262-9541 FAX: 608/262-5974
sanjay@ssec.wisc.edu

Liu Hua-Kuang
Senior Research Engineer/Group Leader
MS 303-310 Jet Propulsion Laboratory
4800 Oak Grove Drive
Pasadena, CA 91109-8099
818/354-8935 FAX: 818/354-9509
hkliu@neuronz.jpl.nasa.gov

Loehr Ed
LASP Space Technology Building
University of Colorado
1234 Innovation Drive
Boulder, CO 80303
303/492-7289 FAX: 303/492-6444

Mankofsky Alan
Science Applications International Corporation
1710 Goodridge Drive
McLean, VA 22102

703/734-5596 FAX: 703/821-1134
alan@mclapo.saic.com

McCollough W. Vance
Hughes Aircraft Company
16800 E. Centertech Parkway
MS A1715
Aurora, CO 80011
303/344-6145 FAX: 303/344-2903

ATTENDEES AISRP WORKSHOP III AUGUST 3-6, 1993

Mihalas **Barbara**
 NCSA/5600 Beckman Institute DR 25
 University of Illinois
 405 N. Mathews Street
 Urbana, IL 61801
 217/244-0637 FAX: 217/244-2909
 bmihalas@ncsa.uiuc.edu

Miller **Glenn**
 Space Telescope Science Institute
 3700 San Martin Drive
 Baltimore, MD 21218
 410/338-4738 FAX: 410/338-1592
 miller@stsci.edu

Mucklow **Glenn**
 Code ST
 Information Systems Branch
 National Aeronautics andSpace Administration
 Washington, DC 20546
 202/358-2235 FAX: 202/358-4166
 gmucklow@leda.hq.nasa.gov

Muller **Ron**
 Mission to Planet Earth Office
 Code 170
 Goddard Space Flight Center
 Greenbelt, MD 20771
 301/286-9695 FAX: 301/286-2477

Pathi **Sridhar**
 Climate System Research Program
 Texas A & M University
 College Station, TX 77843-3150
 409/862-4342 FAX: 409/862-4132
 pathi@csr.p.tamu.edu

Pavlidis **Theo**
 Department of Computer Science
 State University of New York at Stony Brook
 Stony Brook, NY 11794-4400
 516/632-8465 FAX: 516/632-8334
 theo@sbc.suny.edu

Pratt **Terrence W.**
 Code 930.5 CESDIS
 Goddard Space Flight Center
 National Aeronautics andSpace Administration
 Greenbelt, MD 20771
 301/286-4108 FAX: 301/286-5152
 pratt@cesdis1.gsfc.nasa.gov

Raben **Vern**
 Space Environment Laboratory
 NOAA
 325 Broadway
 Boulder, CO 80303

ATTENDEES AISRP WORKSHOP III AUGUST 3-6, 1993

Reichenbach Steve
 Computer Science & Engineering Department
 University of Nebraska
 Lincoln, NE 68588-0115

402/472-5007 FAX: 402/472-7767
 reich@unl.edu

Rotar Paul
 NCAR-SCD
 1850 Table Mesa Road
 Boulder, CO 80306

303/497-1277 FAX: 303/497-1131
 rotar@ncar.ucar.edu

Santek David A.
 Space Science and Engineering Center
 University of Wisconsin-Madison
 1225 W. Dayton Street
 Madison, WI 53706
 608/263-7410
 daves@ssec.wisc.edu

Searight Keith R.
 Department of Atmospheric Sciences
 University of Illinois
 105 S. Gregory
 Urbana, IL 61801
 217/333-8132 FAX: 217/244-4393
 keith@vista.atmos.uiuc.edu

Shaw Crystal
 NCSA/Department of Atmospheric Sciences
 University of Illinois
 405 N. Mathews Street
 Urbana, IL 61801
 217/244-1982 FAX: 217/244-2909
 shaw@ncsa.uiuc.edu

Simmons Karen E.
 Laboratory for Atmospheric and Space Physics
 University of Colorado
 1234 Innovation Drive
 Boulder, CO 80309
 303/492-8363 FAX: 303/493-6444
 BPER::SIMMONS or PISCES::SIMMONS

Singer Howard J.
 Space Environment Laboratory
 NOAA
 325 Broadway
 Boulder, CO 80303

Stetina Fran
 Code 930.8 International Data Systems Office
 Goddard Space Flight Center
 National Aeronautics & Space Administration
 Greenbelt, MD 20771
 301/286-5717 301/286-1635

ATTENDEES AISRP WORKSHOP III AUGUST 3-6, 1993

Storer James A.
 Computer Science Department
 Brandeis University
 Waltham, MA 02254

617/736-2714 FAX: 617/736-2741
 storer@cs.brandeis.edu

Szuszczewicz Edward
 Laboratory for Atmospheric and Space Sciences
 Science Applications International Corporation
 1710 Goodridge Drive
 McLean, VA 22102

703/734-5516 FAX: 703/821-1134
 szusz@mclapo.saic.com

Walsh John E.
 Department of Atmospheric Sciences
 University of Illinois
 105 S. Gregory Avenue
 Urbana, IL 61801

217/333-7521 FAX: 217/244-4393
 walsh@uiatma.atmos.uiuc.edu

Walton Amy
 Science Data Analysis & Computing Systems
 Jet Propulsion Laboratory
 4800 Oak Grove Drive M.S. 180-404
 Pasadena, CA 91109

818/354-3469 FAX: 818/393-4468
 AWalton@nasamail.nasa.gov

Wilhelmson Robert
 NCSA/Department of Atmospheric Sciences
 University of Illinois
 405 N. Mathews Street
 Urbana, IL 61801

217/244-6833 FAX: 217/244-2909
 bw@ncsa.uiuc.edu

Wojtowicz David
 NCSA/Department of Atmospheric Science
 University of Illinois - 5249 Beckman Institute
 405 N. Mathews Street
 Urbana, IL 61801

217/244-1982 FAX: 217/244-2909
 davidw@ncsa.uiuc.edu

Yin John
 M.S. 233-18
 Ames Research Center
 Moffett Field, CA 94035

415/604-3312 415/604-6999
 yin@atlas.arc.nasa.gov

Zonfrelli Joseph V.
 Space Environment Laboratory
 NOAA
 325 Broadway
 Boulder, CO 80303



APPENDIX C

Abstracts



95-82
11/11/95
1-1

**An Interactive Environment for the Analysis of
Large Earth Observation and Model Data Sets**

Principal Investigator: Assistant Professor Kenneth P. Bowman
University of Illinois

Co-Investigators: Professor John E. Walsh
University of Illinois

Professor Robert B. Wilhelmson
University of Illinois

Summary:

We propose to develop an interactive environment for the analysis of large Earth science observation and model data sets. We will use a standard scientific data storage format and a large capacity (>20 GB) optical disk system for data management; develop libraries for coordinate transformation and regridding of data sets; modify the NCSA X Image and X DataSlice software for typical Earth observation data sets by including map transformations and missing data handling; develop analysis tools for common mathematical and statistical operations; integrate the components described above into a system for the analysis and comparison of observations and model results; and distribute software and documentation to the scientific community.

52-82
11/11/87
6.3.1.2
P.1

**Interactive Interface for National Center
for Atmospheric Research (NCAR) Graphics**

Principal Investigator: Dr. William Buzbee
National Center for Atmospheric Research

Co-Investigators: Robert L. Lackman
National Center for Atmospheric Research

Summary:

NCAR Graphics is a FORTRAN 77 library of over 30 high-level graphics modules which are heavily used by science and engineering researchers at over 1500 sites world-wide including many universities and government agencies. These Earth science oriented modules now have a FORTRAN callable subroutine interface which excludes their use by non-programming researchers. This proposal outlines the development of a fully interactive "point and click" menu-based interface using the prevailing toolkit standard for the X-Window System. Options for direct output to the display window and/or output to a Computer Graphics Metafile (CGM) will be provided. X, PEX, and PHIGS will be implemented as the underlying windowing and graphics standards. Associated meteorological and geometric data sets would exploit the network extended NASA Common Data Format, netCDF.

S2-37
11-11-01
1-1

**Development of a Tool-Set for Simultaneous,
Multi-Site Observations of Astronomical Objects**

Principal Investigator: Dr. Supriya Chakrabarti
Boston University

Co-Investigators: Dr. J. Garrett Jernigan
University of California, Berkeley

Dr. Herman L. Marshall
University of California, Berkeley

Summary

A network of ground and space based telescopes can provide continuous observation of astronomical objects. In a "Target of Opportunity" scenario triggered by the system, any telescope on the network may request supporting observations. We propose to develop a set of data collection and display tools to support these observations. We plan to demonstrate the usefulness of this toolset for simultaneous multi-site observations of astronomical targets. Possible candidates for the proposed demonstration include the Extreme Ultraviolet Explorer, International Ultraviolet Explorer, ALEXIS, and sounding rocket experiments. Ground based observations operated by the University of California, Berkeley; the Jet Propulsion Laboratory; and Fairborn Observatory, Mesa, Arizona will be used to demonstrate the proposed concept. Although the demonstration will involve astronomical investigations, these tools will be applicable to a large number of scientific disciplines. The software tools and systems developed as a result of our work will be made available to the scientific community.

34-90

NES 2045

62704

P-1

**Multivariate Statistical Analysis Software Technologies for
Astrophysical Research Involving Large Data Bases**

Principal Investigator: Professor Stanislav Djorgovski
California Institute of Technology

Summary:

The existing and forthcoming data bases from NASA missions contain an abundance of information whose complexity cannot be efficiently tapped with simple statistical techniques. Powerful multivariate statistical methods already exist which can be used to harness much of the richness of these data. Automatic classification techniques have been developed to solve the problem of identifying known types of objects in multiparameter data sets, in addition to leading to the discovery of new physical phenomena and classes of objects. We propose an exploratory study and integration of promising techniques in the development of a general and modular classification/analysis system for very large data bases, which would enhance and optimize data management and the use of human research resources.

A Land-Surface Testbed for EOSDIS**Principal Investigator:**

Dr. William Emery
University of Colorado

Co-Investigators:

Dr. Jeff Dozier
University of California, Santa Barbara

Paul Rotar
National Center for Atmospheric Research

Summary:

We propose to develop an on-line data distribution and interactive display system for the collection, archival, distribution and analysis of operational weather satellite data for applications in land surface studies. A 1,000 km² scene of the western U.S. (centered on the Colorado Rockies) will be extracted from Advanced Very High Resolution Radiometer (AVHRR) imagery collected from morning and afternoon passes of the NOAA polar-orbiters at the direct readout stations operated by CU/CCAR. All five channels of these AVHRR data will be navigated and map registered at CU/CCAR and then be transferred to NCAR for storage in an on-line data system. Software will also be available at NCAR to process and navigate the raw AVHRR data as needed. A display workstation software, based on a Macintosh II computer, will be developed that will display and further process the AVHRR data for studies of vegetation monitoring and snowpack assessment. Various options of presently used techniques for both vegetation and snowpack monitoring will be implemented in the workstation software to provide the individual investigator with the freedom to interact with the satellite image data. The display software will be freely distributed online to interested investigators and the AVHRR data will be made available on-line to anyone interested. In addition, potential users will be sought out and connected to the on-line data archive. This experiment with an active on-line archive and interactive analysis systems will provide experience with a small scale EOSDIS.

56-82

11/15/11-1

11/27/11

P.2

**Geographic Information System for Fusion and Analysis of
High-Resolution Remote Sensing and Ground Truth Data**

Principal Investigator: Anthony Freeman
Jet Propulsion Laboratory

Co-Investigators: Jo Bea Way
Jet Propulsion Laboratory

Pascale Du Bois
Jet Propulsion Laboratory

Franz Leberl
VEXCEL Corporation

Summary:

We seek to combine high-resolution remotely sensed data with models and ground truth measurements, in the context of a Geographical Information System, integrated with specialized image processing software. We will use this integrated system to analyze the data from two Case Studies, one at a boreal forest site, the other a tropical forest site. We will assess the information content of the different components of the data, determine the optimum data combinations to study biogeophysical changes in the forest, assess the best way to visualize the results, and validate the models for the forest response to different radar wavelengths/polarizations.

During the 1990's, unprecedented amounts of high-resolution images from space of the Earth's surface will become available to the applications scientist from the LANDSAT/TM series, European and Japanese ERS-1 satellites, RADARSAT and SIR-C missions. When the Earth Observation Systems (EOS) program is operational, the amount of data available for a particular site can only increase. The interdisciplinary scientist, seeking to use data from various sensors to study his site of interest, may be faced with massive difficulties in manipulating such large data sets, assessing their information content, determining the optimum combinations of data to study a particular parameter, visualizing his results and validating his model of the surface. The techniques to deal with these problems are also needed to support the analysis of data from NASA's current program of Multi-sensor Airborne Campaigns, which will also generate large volumes of data.

In the Case Studies outlined in this proposal, we will have somewhat unique data sets. For the Bonanza Creek Experimental Forest (Case I) calibrated DC-8 SAR data and extensive ground truth measurement are already at our disposal. The data set shows documented evidence to temporal change. The Belize Forest Experiment (Case II) will produce calibrated

DC-8 SAR and AVIRIS data, together with extensive measurements on the tropical rain forest itself. The extreme range of these sites, one an Arctic forest, the other a tropical rain forest, has been deliberately chosen to find common problems which can lead to generalized observations and unique problems with data which raise issues for the EOS System.

57-61

ALC 3091

P-1

**Construction of an Advanced Software Tool
for Planetary Atmospheric Modeling**

Principal Investigator: Dr. Peter Friedland
Ames Research Center

Co-Investigators: Dr. Richard M. Keller
Ames Research Center

Dr. Christopher P. McKay
Ames Research Center

Michael H. Sims
Ames Research Center

Dr. David E. Thompson
Ames Research Center

Summary:

Scientific model-building can be a time intensive and painstaking process, often involving the development of large complex computer programs. Despite the effort involved, scientific models cannot be distributed easily and shared with other scientists. In general, implemented scientific models are complicated, idiosyncratic, and difficult for anyone but the original scientist/programmer to understand. We propose to construct a scientific modeling software tool that serves as an aid to the scientist in developing, using and sharing models. The proposed tool will include an interactive intelligent graphical interface and a high-level domain-specific modeling language. As a testbed for this research, we propose to develop a software prototype in the domain of planetary atmospheric modeling.

System of Experts for Intelligent Data Management (SEIDAM)

Principal Investigator: Dr. David G. Goodenough
Canada Centre for Remote Sensing (CCRS)

Co-Investigators: Joji Iisaka
Canada Centre for Remote Sensing

Ko Fung
University of Ottawa

Summary:

It is proposed to conduct research and development on a system of expert systems for intelligent data management (SEIDAM). CCRS has much expertise in developing systems for integrating geographic information with space and aircraft remote sensing data and in managing large archives of remotely sensed data. SEIDAM will be composed of expert systems grouped in three levels. At the lowest level, the expert systems will manage and integrate data from diverse sources, taking account of symbolic representation differences and varying accuracies. Existing software can be controlled by these expert systems, without rewriting existing software into an Artificial Intelligence (AI) language. At the second level, SEIDAM will take the interpreted data (symbolic and numerical) and combine these with data models. At the top level, SEIDAM will respond to user goals for predictive outcomes given existing data. The SEIDAM Project will address the research areas of expert systems, data management, storage and retrieval, and user access and interfaces.

S. 32
11/10
S. 32
P-1

Knowledge-based Assistance for Science Visualization and Analysis Using Large Distributed Databases

Principal Investigator: Thomas H. Handley, Jr.
Jet Propulsion Laboratory

Co-Investigators: Dr. Allan S. Jacobson
Jet Propulsion Laboratory

Dr. Richard J. Doyle
Jet Propulsion Laboratory

Dr. Donald J. Collins
Jet Propulsion Laboratory

Summary:

Within this decade, the growth in complexity of exploratory data analysis and the sheer volume of space data require new and innovative approaches to support science investigators in achieving their research objectives. To date, there have been numerous efforts addressing the individual issues involved in inter-disciplinary, multi-instrument investigations. However, while successful in small scale, these efforts have not proven to be open and scaleable.

This proposal addresses four areas of significant need: scientific visualization and analysis; science data management; interactions in a distributed, heterogeneous environment; and knowledge-based assistance for these functions. The fundamental innovation embedded within this proposal is the integration of three automation technologies, namely, knowledge-based expert systems, science visualization and science data management. This integration is based on the concept called the DataHub. With the DataHub concept, NASA will be able to apply a more complete solution to all nodes of a distributed system. Both computation nodes and interactive nodes will be able to effectively and efficiently use the data services (access, retrieval, update, etc.) with a distributed, interdisciplinary information system in a uniform and standard way. This will allow the science investigators to concentrate on their scientific endeavors, rather than to involve themselves in the intricate technical details of the systems and tools required to accomplish their work. Thus, science investigators need not be programmers. The emphasis will be on the definition and prototyping of system elements with sufficient detail to enable data analysis and interpretation leading to publishable scientific results. In addition, the proposed work includes all the required end-to-end components and interfaces to demonstrate the completed concept.

Experimenter's Laboratory for Visualized Interactive Science

Principal Investigator: Elaine R. Hansen
University of Colorado at Boulder

Co-Investigators: Marjorie K. Klemp
University of Colorado at Boulder

Sally W. Lasater
University of Colorado at Boulder

Marti R. Szczur
Goddard Space Flight Center

Joseph B. Klemp
National Center for Atmospheric Research

Summary:

The science activities of the 1990's will require the analysis of complex phenomena and large diverse sets of data. In order to meet these needs, we must take advantage of advanced user interaction techniques: modern user interface tools; visualization capabilities; affordable, high performance graphics workstations; and interoperable data standards and translator. To meet these needs, we propose to adopt and upgrade several existing tools and systems to create an experimenter's laboratory for visualized interactive science. Intuitive human-computer interaction techniques have already been developed and demonstrated at the University of Colorado. A Transportable Applications Executive (TAE+), developed at GSFC, is a powerful user interface tool for general purpose applications. A 3D visualization package developed by NCAR provides both color-shaded surface displays and volumetric rendering in either index or true color. The Network Common Data Form (NetCDF) data access library developed by Unidata supports creation, access and sharing of scientific data in a form that is self-describing and network transparent. The combination and enhancement of these packages constitutes a powerful experimenter's laboratory capable of meeting key science needs of the 1990's. This proposal encompasses the work required to build and demonstrate this capability.

211-82
A123.2 N11
62111
p.1

Topography from Shading and Stereo

Principal Investigator:

Professor Berthold P. Horn
Massachusetts Institute of Technology

Co-Investigators:

Michael Caplinger
Arizona State University

Summary:

Methods exploiting photometric information in images that have been developed in machine vision can be applied to planetary imagery. Present techniques, however, focus on one visual cue, such as shading or binocular stereo, and produce results that are either not very accurate in an absolute sense or provide information only at few points on the surface. We plan to integrate shape from shading, binocular stereo and photometric stereo to yield a robust system for recovering detailed surface shape and surface reflectance information. Such a system will be useful in producing quantitative information from the vast volume of imagery being received, as well as in helping visualize the underlying surface. The work will be carried out on a popular computing platform so that it will be easily accessible to other workers.

492-32
1137-1497
P-1

**A Distributed System for Visualizing and Analyzing
Multivariate and Multidisciplinary Data**

Principal Investigator: Dr. Allan S. Jacobson
Jet Propulsion Laboratory

Co-Investigators: Dr. Mark Allen
Dr. Michael Bailey
Dr. Ronald Blom
Leo Blume
Dr. Lee Elson
[all from Jet Propulsion Laboratory]

Summary:

The Linked Windows Interactive Data System (LinkWinds) is being developed with NASA support. The objective of this proposal is to adapt and apply that system in a complex network environment containing elements to be found by scientists working multidisciplinary teams on very large scale and distributed data sets. The proposed three year program will develop specific visualization and analysis tools, to be exercised locally and remotely in the LinkWinds environment, to demonstrate visual data analysis, interdisciplinary data analysis and cooperative and interactive televisualization and analysis of data by geographically separated science teams. These demonstrations will involve at least two science disciplines with the aim of producing publishable results.

S13 82
11/13 11/13/87
11/13
P-1

**The Grid Analysis and Display System (GRADS):
A Practical Tool for Earth Science Visualization**

Principal Investigator:

Dr. James L. Kinter, III
University of Maryland

Summary:

We propose to develop and enhance a workstation based grid analysis and display software system for Earth science dataset browsing, sampling and manipulation. The system will be coupled to a supercomputer in a distributed computing environment for near real-time interaction between scientists and computational results.

214-514
P-1

**Planetary Data Analysis and Display System:
A Version of PC-McIDAS**

Principal Investigator: Dr. Sanjay S. Limaye
University of Wisconsin-Madison

Co-Investigators: L. A. Sromovsky
University of Wisconsin-Madison

R. S. Saunders
Jet Propulsion Laboratory

Michael Martin
Jet Propulsion Laboratory

Summary:

We propose to develop a system for access and analysis of planetary data from past and future space missions based on an existing system, the PC-McIDAS workstation. This system is now in use in the atmospheric science community for access to meteorological satellite and conventional weather data. The proposed system would be usable by not only planetary atmospheric researchers but also by the planetary geologic community. By providing the critical tools of an efficient system architecture, newer applications and customized user interfaces can be added by the end user within such a system.

P-1

**Multi-Layer Holographic Bifurcative Neural Network
System for Real-Time Adaptive EOS Data Analysis**

Principal Investigator: Dr. Hua-Kuang Liu
Jet Propulsion Laboratory

Co-Investigators: Professor K. Huang
University of Southern California

J. Diep
Jet Propulsion Laboratory

Summary:

Optical data processing techniques have the inherent advantage of high data throughput, low weight and low power requirements. These features are particularly desirable for onboard spacecraft in-situ real-time data analysis and data compression applications. The proposed multi-layer optical holographic neural net pattern recognition technique will utilize the nonlinear photorefractive devices for real-time adaptive learning to classify input data content and recognize unexpected features. Information can be stored either in analog or digital form in a nonlinear photorefractive device. The recording can be accomplished in time scales ranging from milliseconds to microseconds. When a system consisting of these devices is organized in a multi-layer structure, a feedforward neural net with bifurcating data classification capability is formed. The interdisciplinary research will involve the collaboration with top digital computer architecture experts at the University of Southern California.

216-82
112. 11. 11
1.1

Development of an Expert Data Reduction Assistant

Principal Investigator: Dr. Glenn E. Miller
Space Telescope Science Institute

Co-Investigators: Dr. Mark D. Johnston
Space Telescope Science Institute

Dr. Robert J. Hanisch
Space Telescope Science Institute

Summary:

We propose the development of an expert system tool for the management and reduction of complex data sets. The proposed work is an extension of a successful prototype system for the calibration of CCD images developed by Dr. Johnston in 1987. (ref.: Proceedings of the Goddard Conference on Space Applications of Artificial Intelligence)

The reduction of complex multi-parameter data sets presents severe challenges to a scientist. Not only must a particular data analysis system be mastered, (e.g. IRAF/SDAS/MIDAS), large amounts of data can require many days of tedious work and supervision by the scientist for even the most straightforward reductions. The proposed Expert Data Reduction Assistant will help the scientist overcome these obstacles by developing a reduction plan based on the data at hand and producing a script for the reduction of the data in a target common language.

517-92
NASA ONLY

**VIEWCACHE: An Incremental Pointer-based Access
Method for Autonomous Interoperable Databases**

Principal Investigator: Associate Professor N. Roussopoulos
University of Maryland

Co-Investigators: Dr. Timos Sellis
University of Maryland

Summary:

One of biggest problems facing NASA today is to provide scientists efficient access to a large number of distributed databases. Our pointer-based incremental database access method, VIEWCACHE, provides such an interface for accessing distributed datasets and directories. VIEWCACHE allows database browsing and search performing inter-database cross-referencing with no actual data movement between database sites. This organization and processing is especially suitable for managing Astrophysics databases which are physically distributed all over the world. Once the search is complete, the set of collected pointers pointing to the desired data are cached. VIEWCACHE includes spatial access methods for accessing image datasets, which provide much easier query formulation by referring directly to the image and very efficient search for objects contained within a two-dimensional window. We will develop and optimize a VIEWCACHE External Gateway Access to database management systems to facilitate distributed database search.

513-22
11/23/07
11

**Advanced Data Visualization and Sensor Fusion:
Conversion of Techniques from Medical Imaging to Earth Science**

Principal Investigator: Dr Richard C. Savage
Hughes Aircraft Company

Co-Investigators: Dr. Chin-Tu Chen
University of Chicago

Dr. Charles Pelizzari
University of Chicago

Dr. Veerabhadran Ramanathan
University of Chicago

Summary:

Hughes Aircraft Company and the University of Chicago propose to transfer existing medical imaging registration algorithms to the area of multi-sensor data fusion. The University of Chicago's algorithms have been successfully demonstrated to provide pixel by pixel comparison capability for medical sensors with different characteristics. The research will attempt to fuse GOES, AVHRR, and SSM/I sensor data which will benefit a wide range of researchers.

The algorithms will utilize data visualization and algorithm development tools created by Hughes in its EOSDIS prototyping. This will maximize the work on the fusion algorithms since support software (e.g. input/output routines) will already exist. The research will produce a portable software library with documentation for use by other researchers.

32
10/1/95
P-1

High Performance Compression of Science Data

Principal Investigator: Dr. James A. Storer
Brandeis University

Co-Investigators: Dr. Martin Cohn
Brandeis University

Summary:

In the future, NASA expects to gather over a tera-byte per day of data requiring space for levels of archival storage. Data compression will be a key component in systems that store this data (e.g., optical disk and tape) as well as in communications systems (both between space and Earth and between scientific locations on Earth). We propose to develop algorithms that can be a basis for software and hardware systems that compress a wide variety of scientific data with different criteria for fidelity/bandwidth tradeoffs. The algorithmic approaches we consider are specially targeted for parallel computation where data rates of over 1 billion bits per second are achievable with current technology.

SAVS: A Space Analysis and Visualization System

Principal Investigator: Dr. Edward P. Szuszczewicz
Science Applications International Corporation

Co-Investigators: Dr. Alan Mankofsky
Science Applications International Corporation

Dr. Charles C. Goodrich
University of Maryland

Summary:

We propose to develop, test, demonstrate, and deliver to NASA a powerful and versatile data acquisition, manipulation, analysis and visualization system which will enhance scientific capabilities in the display and interpretation of diverse and distributed data within an integrated user-friendly environment. Our approach exploits existing technologies and combines three major elements into an easy-to-use interactive package: 1) innovative visualization software, 2) advanced database techniques, and 3) a rich set of mathematical and image processing tools. Visualization capabilities will include one-, two-, and three-dimensional displays, along with animation, compression, warping and slicing functions. Analysis tools will include generic mathematical and statistical techniques along with the ability to use large scale models for interactive interpretation of large volume data sets. Our system will be implemented on Sun and DEC UNIX workstations and on the Stardent Graphics Supercomputer. Our final deliverable will include complete documentation and a NASA/NSF-CDAW/SUNDIAL campaign demonstration.

22
111
12
P-1

**A Spatial Analysis and Modeling System (SAMS)
for Environment Management**

Co-Investigators:

Fran Stetina
Goddard Space Flight Center

Dr. John Hill
Louisiana State University

Dr. Paul Chan
Science Systems and Applications, Inc.

Robert Jaske
Federal Emergency Management Agency

Gilbert Rochon
Dillard University

Summary:

This is a proposal to develop a uniform global environmental data gathering and distribution system to support the calibration and validation of remotely sensed data. SAMS is based on an enhanced version of FEMA's Integrated Emergency Management Information Systems and the Department of Defense's Air Land Battlefield Environment Software Systems. This system consists of state-of-the-art graphics and visualization techniques, simulation models, database management and expert systems for conducting environmental and disaster preparedness studies. This software package will be integrated into various Landsat and UNEP-GRID stations which are planned to become direct readout stations during the EOS timeframe. This system would be implemented as a pilot program to support the Tropical Rainfall Measuring Mission (TRMM). This will be a joint NASA-FEMA-University-Industry project.

122-32
27-00**A Distributed Analysis and Visualization
System for Model and Observational Data**

P1

Principal Investigator: Professor Robert Wilhelmson
University of Illinois

Co-Investigators: Dr. Steven Koch
Goddard Space Flight Center

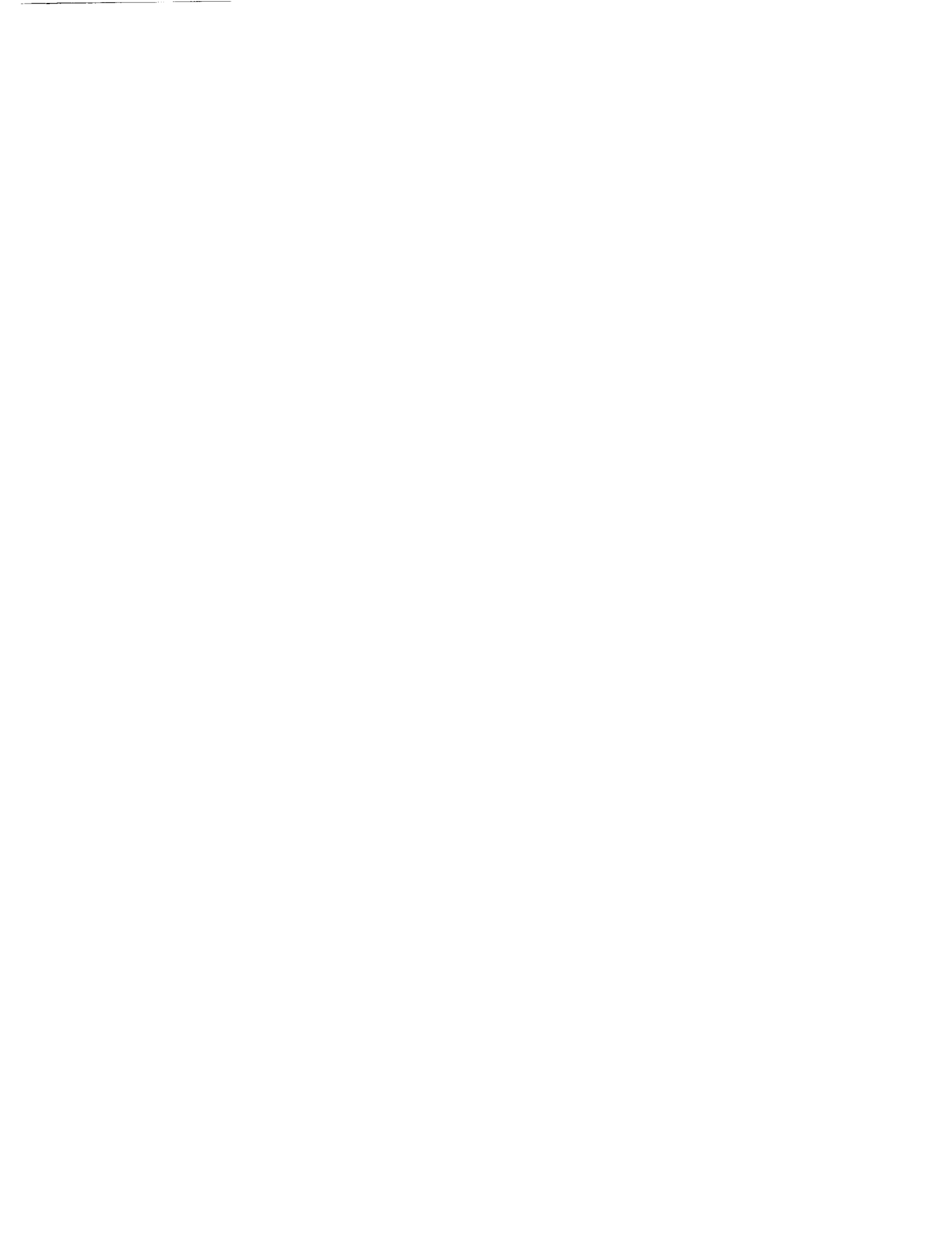
Summary:

The objective of this proposal is to develop an integrated and distributed analysis and display software system which can be applied to all areas of the Earth System Science to study numerical model and earth observational data from storm to global scale. This system will be designed to be easy to use, portable, flexible and easily extensible and to adhere to current and emerging standards whenever possible. It will provide an environment for visualization of the massive amounts of data generated from satellites and other observational field measurements and from model simulations during or after their execution. Two- and three-dimensional animation will also be provided. This system will be based on a widely used software package from NASA called GEMPAK and prototype software for three-dimensional interactive displays built at NCSA. The underlying foundation of the system will be a set of software libraries which can be distributed across a UNIX based supercomputer and workstations.

Handwritten notes:
10/10
10/10
P. 503

APPENDIX D

Demonstrations



Demonstration Schedule for AISRP Workshop

SGI Indigo 2 Extreme sgidemo2	PATHFINDER LinkWinds Datahub	3 - 3:40 3:40 - 4:20 4:20 - 5
SGI Indigo 2 Extreme (Limaye) sgidemo	McIdas-eXplorer	3 - 5
DECstation nowhere	EOS testbed Interactive NCAR Graphics	3 - 4 4 - 5
IBM RS6000 rsdemo	Envision SAVS (Space Data Analysis and Visualization System)	3 - 4 4 - 5
SGI Indigo XS24 picasso2	PolyPaint+ GRADS (Grid Analysis and Display System)	3 - 4 4 - 5
Sun 3 (use as X display for sirius Sparcstation)	SIGMA (Scientists' Intelligent Graphical Modeling Assistant) Software Support Laboratory	3 - 4 4 - 5
VAXstation pup	GGGS (Galileo Geometry and Graphics Software)	3 - 5
MacIntosh	Macsigma0	4 - 5

PATHFINDER

A Distributed System for the Visualization and Analysis of
Observed and Modeled Meteorological Data
PI: Dr. Robert Wilhelmson, Univ. of Illinois

LinkWinds

LinkWinds: The Linked Windows Interactive Data System
PI: Dr. Allan S. Jacobson, JPL

Datahub

DataHub: Knowledge-Based Science Data Management
PI: Tom Handley, JPL

McIdas-eXplorer

Planetary Data Analysis and Display System: A Version of PC-
McIDAS

PI: Dr. Sanjay S. Limaye, University of Wisconsin-Madison

EOS Testbed

A Land-Surface Testbed for the EOS Data Information System
(EOSDIS)

PI: Dr. William Emery, Colorado Center for Astrodynamic
Research

Interactive NCAR Graphics

Interactive NCAR Graphics
PI: Bob Lackman, NCAR

Envision

Envision: An Analysis and Display System for Large

Geophysical Data Sets

PI: Dr. Kenneth Bowman, Texas A&M

SAVS

SAVS: A Space Data Analysis and Visualization System

PI: Dr. E. Szuszczewicz, Laboratory for Atmospheric and Space
Science/SAIC

PolyPaint+

Experimenter's Laboratory for Visualized Interactive Science
PI: Elaine Hansen, University of Colorado

GRADS

The Grid Analysis and Display System (GrADS)

PI: Dr. James L. Kinter, Center for Ocean-Land-Atmosphere
Interactions
University of Maryland

MacSigma0

Geographic Information System (GIS) for Fusion and Analysis
of High-Resolution Remote Sensing and Ground Truth Data
PI: Dr. A. Freeman, JPL

Software Support Laboratory

PI: Randy Davis, University of Colorado/LASP

GGGS

Galileo Geometry and Graphics Software

Karen Simmons, Kirk Benell, University of Colorado/LASP

SIGMA

Scientists' Intelligent Graphical Modeling Assistant

PI: Dr. Richard Keller, Ames Research Center



Demonstration Schedule for AISRP Workshop

SGI Indigo 2 Extreme sgidemo2	PATHFINDER LinkWinds Datahub	3 - 3:40 3:40 - 4:20 4:20 - 5
SGI Indigo 2 Extreme (Limaye) sgidemo	McIdas-eXplorer	3 - 5
DECstation nowhere	EOS testbed Interactive NCAR Graphics	3 - 4 4 - 5
IBM RS6000 rsdemo	Envision SAVS (Space Data Analysis and Visualization System)	3 - 4 4 - 5
SGI Indigo XS24 picasso2	PolyPaint+ GRADS (Grid Analysis and Display System)	3 - 4 4 - 5
Sun 3 (use as X display for sirius Sparcstation)	SIGMA (Scientists' Intelligent Graphical Modeling Assistant) Software Support Laboratory	3 - 4 4 - 5
VAXstation pup	GGGS (Galileo Geometry and Graphics Software)	3 - 5
MacIntosh	Macsigma0	4 - 5

PATHFINDER

A Distributed System for the Visualization and Analysis of
Observed and Modeled Meteorological Data
PI: Dr. Robert Wilhelmson, Univ. of Illinois

LinkWinds

LinkWinds: The Linked Windows Interactive Data System
PI: Dr. Allan S. Jacobson, JPL

Datahub

DataHub: Knowledge-Based Science Data Management
PI: Tom Handley, JPL

McIdas-eXplorer

Planetary Data Analysis and Display System: A Version of PC-
McIDAS

PI: Dr. Sanjay S. Limaye, University of Wisconsin-Madison

EOS Testbed

A Land-Surface Testbed for the EOS Data Information System
(EOSDIS)

PI: Dr. William Emery, Colorado Center for Astrodynamics
Research

Interactive NCAR Graphics

Interactive NCAR Graphics
PI: Bob Lackman, NCAR

Envision

Envision: An Analysis and Display System for Large

Geophysical Data Sets

PI: Dr. Kenneth Bowman, Texas A&M

SAVS

SAVS: A Space Data Analysis and Visualization System
PI: Dr. E. Szuszczewicz, Laboratory for Atmospheric and Space
Science/SAIC

PolyPaint+

Experimenter's Laboratory for Visualized Interactive Science
PI: Elaine Hansen, University of Colorado

GRADS

The Grid Analysis and Display System (GrADS)

PI: Dr. James L. Kinter, Center for Ocean-Land-Atmosphere
Interactions
University of Maryland

MacSigma0

Geographic Information System (GIS) for Fusion and Analysis
of High-Resolution Remote Sensing and Ground Truth Data
PI: Dr. A. Freeman, JPL

Software Support Laboratory

PI: Randy Davis, University of Colorado/LASP

GGGS

Galileo Geometry and Graphics Software

Karen Simmons, Kirk Benell, University of Colorado/LASP

SIGMA

Scientists' Intelligent Graphical Modeling Assistant

PI: Dr. Richard Keller, Ames Research Center



APPENDIX E

Presentation Material

AN OVERVIEW OF THE SOFTWARE SUPPORT LABORATORY

Mr. Randal Davis
Laboratory for Atmospheric and Space Physics
University of Colorado at Boulder

August 3, 1993

LASP

An Overview of the Software Support Laboratory

Presented at the
Applied Information Systems Research Program Workshop III
Boulder, Colorado
3 August 1993

By
Randal Davis
Laboratory for Atmospheric and Space Physics
University of Colorado at Boulder

LASP

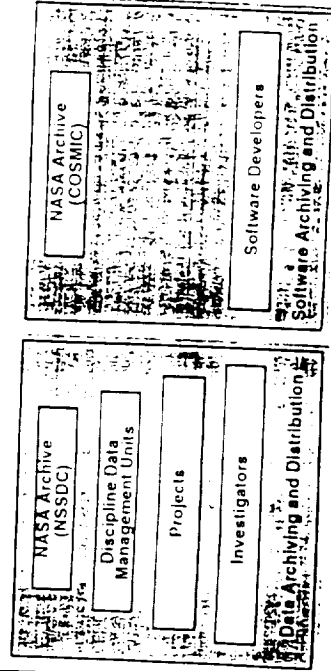
What is the Software Support Laboratory?

- A Research Project
 - Trying to improve the development, archiving and distribution of scientific software
- A Working Software Repository and Distribution Center
 - For software developed under NASA's Applied Information Systems Research Program
- A Source of Information
 - Conveying information on NASA's needs and scientist's desires for software to developers
 - Providing scientists with information on available software products and on software trends

8-0010 11-Aug-93

LASP

Software Archiving and Distribution Is Still Primitive



8-0010 11-Aug-93

LASP

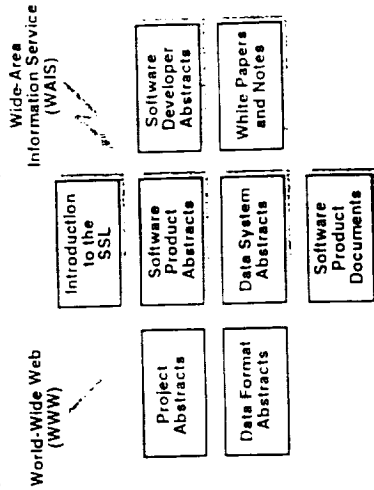
Types of Support for Software Products

- Another Party Maintains and Distributes Software and Documentation
 - SSL provides only an abstract describing the software product
 - Users are directed to the other party for further information or for access to the software
- SSL Maintains and Distributes Software and Documentation
 - Minimal Support
 - Documentation and software available from the SSL in the form provided by the developer
 - Full Support
 - Peer review and usability testing performed on the software
 - Expertise on the software maintained at SSL
 - Porting and customization available
- Integrated Support
 - The party developing or maintaining the software is seamlessly linked with the SSL's information and order services

8 0004 14-0931

LASP

The SSL Online Information Service for Software Developers and Users



8 0004 14-0931

LASP

Earth and Space Science Data CD-ROM Samplers

- A set of CD-ROM disks is being produced by the SSL and CSAT with sample Earth and space science data for use in testing and evaluating software
 - Most NASA science disciplines will be represented:
 - Astrophysics and Astronomy
 - Earth Atmosphere, Ocean and Land
 - Microgravity
 - Planetary
 - Solar-Terrestrial Physics
 - Most common data object types will be included:
 - Images, maps, tables, spectra, and more
 - Most common data formats will be covered:
 - CDF, HDF, NetCDF, PDS, SFDU, and more
 - All datasets will have documentation and software for display

8 0004 14-0931

LASP

Software Sampler

- We are thinking about producing a CD-ROM sampler disk with software for NASA Earth and space scientists:
 - Software developed under the AISRP
 - Other popular application programs
 - Libraries of routines to support access to popular data formats
 - CDF
 - NetCDF
 - PDS
 - SFDU
 - With documentation, installation directions and test data

8 0004 14-0931



What The SSL Would Like From You

- Comments and suggestions regarding the initial 0.1 Alpha version of the SSL Online Information service
- Information for the abstracts on your projects and software
- Decisions on whether or not the SSL will archive and distribute your software
- Comments and suggestions regarding the data to be placed on the Earth and space science sampler CD-ROMs
- Your thoughts on a software sampler CD-ROM

Summary of Data for Sampler CD-ROM Set

Data Set	Project	Data Type	Format	# of Files	Size (MB)				
Planetary — Mars Surface	Viking 1 and 2	Raw Images	PDS	73	34				
Raw Images — Compressed									
Raw Images — Browse									
Mars Digital Image Maps (MDIMs)									
Color MDIMs									
Mars Digital Terrain Model									
Airbrush maps									
Total				188	101				
Planetary — Venus Surface	Magellan	Mosaic Images	PDS/VICAR	80	40				
Compressed Mosaic Images									
Global Emissivity									
Global Reflectivity									
Global Slope									
Global Topography									
Altimetry/Radiometry									
Total							105	65	
Astronomy — Infrared		IRAS				Infrared Sky Survey (Part of North Hemisphere)	FITS	62	43
Infrared Sky Survey (Part of North Hemisphere)									
Astronomy — Ultraviolet	IUE	Spectra of Bright Blue (F Cygn) Stars	FITS	200	4				
Spectra of Bright Blue (F Cygn) Stars									
Astronomy — X-Ray	Einstein	IPC Slow Sky Survey (Part of North Hemisphere)	FITS	124	63				
IPC Slow Sky Survey (Part of North Hemisphere)									
Astronomy — Solar	Skylab	Skylab solar corona images	Special	5	1				
Skylab solar corona images									

The Universal Resource Locator
of the Software Support Laboratory
online information service is:

http://ssllab.colorado.edu:2222/ssl/ssl_homepage.html

See Appendix Page E-48 for NCSA Mosaic

Summary of Data for Sampler CD-ROM Set

Solar-Terrestrial Physics	IE	Spectra Image	SFDU SFDUCDF	4 19 23	14 11 25
Earth Atmosphere	ERIS	Guided Data	COF	5	120
Earth Surface	AVIRIS TIROS CLES	Image (MS) Image (MS) Image (MS)	Special Raw TM Raw TM	2 20 48 70	288 40 96 424
Earth Ocean	DMSP	Image (Map)	HDF	120	12
Microgravity	Spacelab	Image	TIFF	10	2

Holographic Neural Networks for Space and Commercial Applications

APPLIED INFORMATION SYSTEMS
RESEARCH PROGRAM WORKSHOP III

by

DR. HUA-KUANG LIU
JET PROPULSION LABORATORY
CALIFORNIA INSTITUTE OF TECHNOLOGY
PASADENA, CA 91109-8099

BOULDER, COLORADO
3-6 August 1993

HOLOGRAPHIC NEURAL NETWORKS FOR SPACE AND COMMERCIAL APPLICATIONS

Dr. Hua-Kuang Liu
Jet Propulsion Laboratory

August 3, 1993

HOLOGRAPHIC NEURAL NETWORKS

APPROACH

- Develop a system in which a large number of image patterns can be stored and recognition of weak signals can be achieved through amplification in a photorefractive crystal.
- Realize modified signed digit algorithm by optical means
- Apply a family of reference waves with different incident angles to a photorefractive crystal to generate holographically created arrays of laser beams.

HOLOGRAPHIC NEURAL NETWORKS

OBJECTIVE/CONCEPT

- Multi-channel large memory and retrieval capability
- Extremely-weak signal detection for cluttered optical pattern recognition
- Optical computing with high speed operation
- Ultra-fast reconfigurable laser beam array generation for interconnection in computers and neural networks

HOLOGRAPHIC NEURAL NETWORKS

Technical Basis

- Holographic Parallel Template Matching
- Photorefractive Coupling Energy Transfer
 - Computer Generated Holograms
 - Spatial Light Modulation

HOLOGRAPHIC NEURAL NETWORKS

EXAMPLE

- The pattern recognizer can greatly amplify the target signal while filtering out the background noise/clutter.
- Compare to other correlators, the signal-to-noise ratio is *one million times* stronger!
- Partial input of very dimly illuminated targets can be recognized.
- Multi-channel large-capacity neural feedback by electronics can be applied for commercial areas such as finger prints identification.

HOLOGRAPHIC NEURAL NETWORKS

APPLICATIONS

- The neural associative memory and retrieval capability of the MOFENAR has the following applications:
 - (a) NASA applications
 - 1) Robotics vision for Mars rover
 - 2) Automated space station service
 - 3) Space structure health monitoring
 - 4) Data compression
 - 5) Data classification
 - 6) Information storage.

HOLOGRAPHIC NEURAL NETWORKS

APPLICATIONS

- (b) Commercial applications
 - 1) Assembly line quality control
 - 2) Machine vision for industry automation
 - 3) Autonomous navigation of ships and automobiles
 - 4) Collision avoidance of motor vehicles;
 - 5) Automatic traffic control;
 - 6) Advanced security lock;
 - 7) Document authentication
 - 8) Signature recognition
 - 9) Passport identification;
 - 10) Finger print recognition for Department of Motor Vehicles for driver's license validation, and for FBI and Police Department for fast criminal identification.

HOLOGRAPHIC NEURAL NETWORKS
APPLICATIONS

- The real-time 3-D display capability of holography has the following applications:
- (a) NASA Applications
 - 1) Virtual environment for astronaut training
 - 2) Space shuttle operation simulation
 - 3) Space station condition simulation
 - 4) 3-D Virtual reality of Moon and Mars

HOLOGRAPHIC NEURAL NETWORKS
APPLICATIONS

(b) Commercial Applications

- 1) Training of company employees and factory workers to increase their working efficiency and productivity
- 2) The training of policemen in their encounter of criminals
- 3) Compact 3-D commercial advertisement
- 4) 3-D motion picture displays in large stadiums
- 5) Virtual environments for recreation and entertainment
- 6) 3D multi-media communication
- 7) 3-D knowledge epresentation and information integration for homes and offices
- 8) Information storage and display for telecommunication and teleconferencing.

HOLOGRAPHIC NEURAL NETWORKS

COMMERCIALIZATION

- Experimental results showed promising results of the proposed work.
- Private Industries have expressed interests in the new technology and end product of the project.. Examples are Hamamatsu of Japan and Standard Packaging of California.
- JPL has a specific Commercialization Office that has the mission of transferring space technology to private industry.
- Initial commercial invesment of \$25k on 3-D display just arrived in Caltech in July ; \$250k forthcoming in November, 1993.

HOLOGRAPHIC NEURAL NETWORKS

READINESS

- FEASIBILITY Completed '93
- PROOF-OF-CONCEPT '93
- PROTOTYPE '95
- USER'S HANDS 2000

FUNDING SUPPORT

	FY \$k		
	'91	'92	'93 '94
ARPA	75	160	160 95
NASA	120	126	134 80?

THE SEIDAM (SYSTEM OF EXPERTS FOR INTELLIGENT DATA MANAGEMENT) PROJECT

**Mr. David G. Goodenough
Mr. Ko B. Fung**
Department of Natural Resources Canada

August 3, 1993

THE SEIDAM (SYSTEM OF EXPERTS FOR INTELLIGENT DATA MANAGEMENT) PROJECT

David G. Goodenough and Ko B. Fung
Department of Natural Resources Canada

**AISR Program
August 3, 1993**

©1993

SEIDAM

OUTLINE

- ◆ The need for intelligent data management.
- ◆ The SEIDAM Concept
- ◆ Three test sites
- ◆ SEIDAM software and future activities
- ◆ The SEIDAM outputs and summary

2

SEIDAM

THE NEED FOR INTELLIGENT DATA MANAGEMENT

- ◆ Monitoring a substantial area requires the integration of data and knowledge from many sources.
- ◆ The U.S.A. and Canada will receive more than 1 terrabyte per day by the year 2000.
- ◆ New remote sensing sensors can give canopy chemistry and direct measurements of timber volume not possible before.
- ◆ Intelligent data management systems are required to reduce complexity, increase adaptability, reduce costs, enhance cooperation amongst distributed data and knowledge sources, to improve the dissemination of information, and to provide an audit trail of data and information.

3

SEIDAM

A RESPONSE

- ◆ Under development is the System of Experts for Intelligent Data Management (SEIDAM) (1991-1996).
- ◆ It is expected that data distribution policies will favor much more access to data. SEIDAM is a NASA project.
- ◆ SEIDAM is also supported by the Government of Canada and the Province of British Columbia.
- ◆ The SEIDAM project (phase 2) will complete in 1996.
- ◆ The SEIDAM software contains expert systems to provide answers.
- ◆ SEIDAM learns from examples or cases.
- ◆ A case consists of a query (question), the data, and the desired goal.

SEIDAM

EXAMPLES OF QUERIES

A study has been conducted with the provincial partners in order to determine the best set of queries with which to implement SEIDAM.

- ◆ How much forest do (or did or will) we have in this test site?"
- ◆ Do - update the existing forest inventory with the most recent remote sensing imagery.
- ◆ Did - use multitemporal imagery and GIS files to build up a temporal description of the forests over time.
- ◆ Will - use models to predict forest growth and pest spread under changing climate conditions.

SEIDAM

MORE QUERIES

- ◆ What is the present timber volume of Douglas fir in this watershed?
- ◆ "Who were the sources of the data used to provide this answer?"
- ◆ What is the yearly rate of forest depletion in this test site over the past 20 years?
- ◆ Where has bud flush (new growth) been reduced by damage agents?

SEIDAM

PRODUCT GOALS

An alternative to a query is a product. A product may answer many queries. Examples of product goals are:

- maps,
- updated GIS files,
- tabular summaries,
- visualizations,
- temporal change products,
- text, values,
- history of processing
- characteristics and accuracies of the products
- previous dissemination information

SEIDAM

SYSTEM OF EXPERTS FOR INTELLIGENT DATA MANAGEMENT (SEIDAM)

- PRINCIPAL INVESTIGATOR: DAVID GOODENOUGH
- CO-INVESTIGATOR (CCRS): KO FUNG
- CO-INVESTIGATOR (CCRS): DIANE RICHARDSON
- CO-INVESTIGATOR (CCRS): KARL STAENZ
- CO-INVESTIGATOR (PFC): ALAN THOMSON
- CO-INVESTIGATOR (BCMOPF): JOHN WAKELIN
- CO-INVESTIGATOR (BCMELP): ROGER BALSER
- CO-INVESTIGATOR (JRC-I): ALOIS SIEBER
- CO-INVESTIGATOR (KTH): KENNERT TORLEGARD
- CO-INVESTIGATOR (KTH): HANS HAUSKA

NASA'S APPLIED INFORMATION SYSTEMS RESEARCH PROGRAM.

SEIDAM

ANSWERING A QUERY

- ◆ How do we present the answers to the query?
 - ◆ We will need an output translator which puts the answer into a format that the user can easily assimilate.
 - ◆ Three components must be addressed:
 - mode of information presentation,
 - task environment of the decision,
 - individual characteristics of the decision maker.
 - ◆ How much visualization will be needed?
- In addition, we need to address
- ◆ How to transmit the information to others
 - ◆ How to describe and present the processing history

SEIDAM

WHICH AGENCIES ARE INVOLVED IN SEIDAM?

- ENERGY, MINES AND RESOURCES CANADA
 - CANADA CENTRE FOR REMOTE SENSING
- FORESTRY CANADA
 - PACIFIC FORESTRY CENTRE
- U.S. NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
 - APPLIED INFORMATION SYSTEMS RESEARCH PROGRAM
- INDUSTRY, SCIENCE AND TECHNOLOGY CANADA
 - STRATEGIC TECHNOLOGIES BRANCH
- B.C. MINISTRY OF FORESTS
 - INVENTORY BRANCH
- B.C. MINISTRY OF ENVIRONMENT, LANDS AND PARKS
 - SURVEYS AND RESOURCE MAPPING BRANCH
- EEC JOINT RESEARCH CENTRE AT ISPRA, ITALY
 - MICROWAVE SIGNATURES LABORATORY
- ROYAL INSTITUTE OF TECHNOLOGY (KTH)

SEIDAM

SEIDAM OBJECTIVE

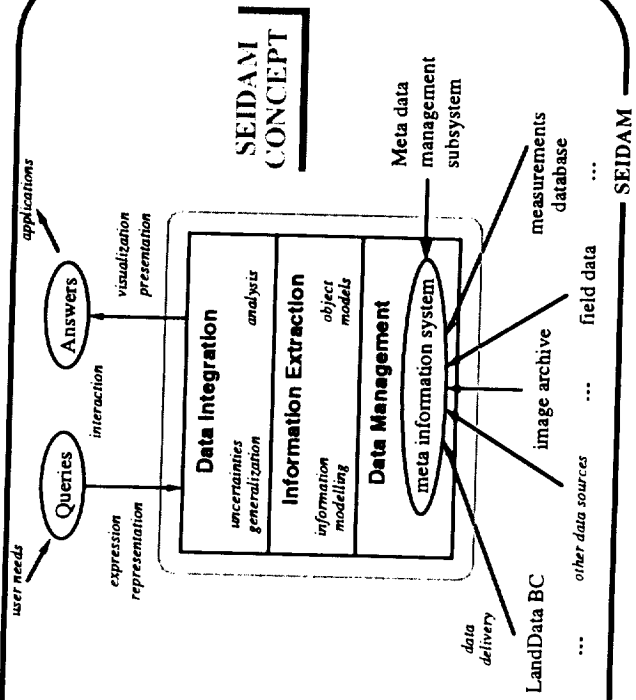
To create a system of experts for intelligent data management (SEIDAM) which will integrate remote sensing data from satellites and aircraft with geographic information systems and manage large archives of remotely sensed data for dynamic selection of data sources and sensor characteristics for query-based recognition of forest objects appropriate for environmental forest monitoring.

SEIDAM

Meta Data

- ◆ Raw and Derived Image Products
- ◆ GIS products
- ◆ Project Information
- ◆ Science Models
- ◆ Publications
- ◆ Field Data
- ◆ Project Standards

SEIDAM CONCEPT

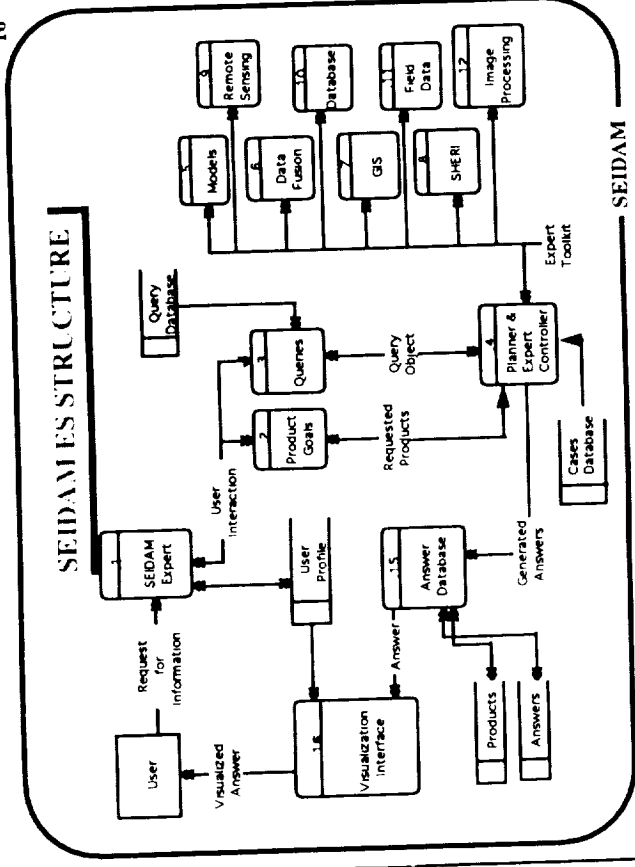
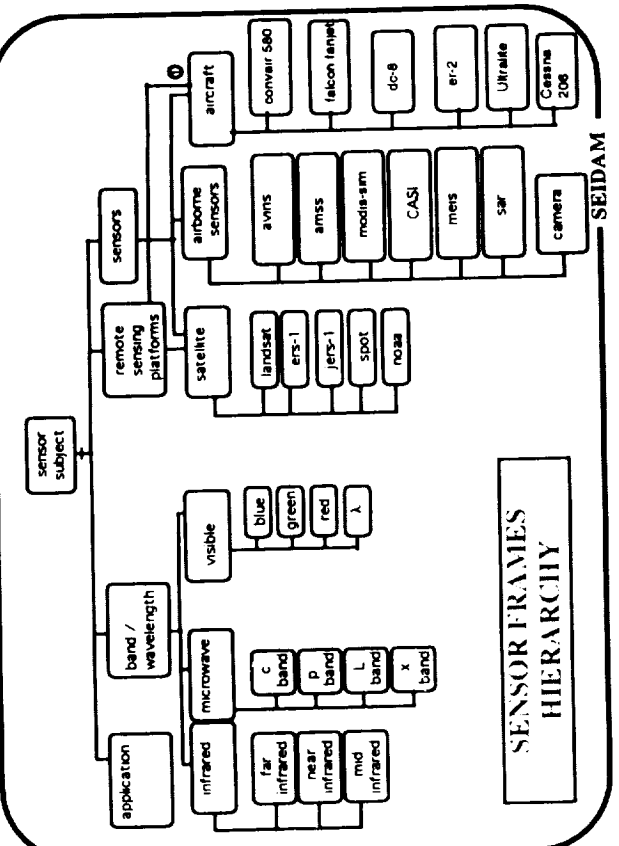


Meta Data Subsystem

- ◆ meta data creation, insertion, update and transfer
- ◆ meta data presentation
- ◆ automatic data insertion and update
- ◆ data security
- ◆ linkage to the archival
- ◆ query handling

Information Tracking

- ◆ distribution history
- ◆ Information about existence
- ◆ accuracy
- ◆ transmission
- ◆ file format standard for meta data exchange
- ◆ algorithms and software used to generate the related data products
- ◆ organization and personnel involved
- ◆ Assumptions
- ◆ Parameters used
- ◆ Instrument



PROBLEMS WITH MULTIPLE DATA SOURCES

- ◆ Costs of acquisition
- ◆ It is more complex to handle multiple sensors.
- ◆ Different sensors may have differing viewing geometries
- ◆ Different sensors may have differing responses to the atmosphere.
- ◆ Different sensors may have differing recording formats.
- ◆ Generalization may be needed to handle different scales.

WHY USE MULTIPLE DATA SOURCES?

- ◆ A single data source may miss significant characteristics needed to identify an object.
- ◆ A single data source may be subjected to noise.
- ◆ Multiple data sources provide complementary information.
- ◆ Redundant information increases the accuracy and certainty of object identification.
- ◆ One is more likely to have data available at the desired time.
- ◆ Historical GIS information can guide recognition process.

SPATIAL ASPECTS FOR GIS & RS

- ◆ What is the reference datum to be used eg. NAD'27 or NAD'83?
- ◆ Generalizations and abstractions used for base map may create spatial displacements with respect to image data.
- ◆ Different GIS sources may have spatial errors between their respective base maps
- ◆ Is a DEM to be used to geometrically correct the imagery?
- ◆ What is the accuracy of the DTM in meters and is this figure greater than the maximum potential image shift due to terrain? Has the DTM been generalized? B.C. TRIM data is ungeneralized.

SEIDAM

MISMATCHES BETWEEN GIS DATA AND REMOTE SENSING DATA

- ◆ Sensor resolution may not be sufficient to record spatial features used for polygon delineation.
- ◆ GIS features may not be visible in images due to non-stationarity of objects (e.g. Crops, water boundaries, burns, etc.).
- ◆ Image objects reflect seasonal variations.

SEIDAM

DATA SOURCE SELECTION

- ◆ Select data source with best coverage of the ground area..
- ◆ Select data source with lowest inherent errors such as geometric and sensor errors.
- ◆ Select data source with optimum spatial resolution.
 - We don't want too coarse a resolution so that desired objects can not be identified.
 - Also, we don't want too fine a spatial resolution so that desired objects become too complex in the image.

SEIDAM

DATA SOURCE SELECTION #2

- ◆ Select best wavelength intervals of the sensor measurements for the desired objects.
- ◆ Select data sources for which there exist implemented algorithms.
- ◆ Select data source with lowest cost pre-processing required.
- ◆ Select data source with the least computational cost of the analysis to derive the required attributes.

SEIDAM

TRIM

- ◆ The Terrain Resource Information Management (TRIM) program provides a DTM for British Columbia at a presentation scale of 1:20,000.
- ◆ TRIM data conform to the reference datum NAD'83 and adhere to stringent positional accuracy rules.
- ◆ In SEIDAM, TRIM data provide the common georeferencing base for other data sets, particularly satellite imagery, and GPS derived data, becomes much easier.

SEIDAM

SEIDAM Data Sources

Satellite:

LANDSAT 5,6
SPOT 3
ERS-1
JERS-1
NOAA/AVHRR

6 SATELLITE SENSORS

Aircraft:

CCRS SAR (X, C - polarimetric)
CCRS MEIS (push broom scanner)
CCRS AMSS
NASA airborne SAR (C, L, P - polarimetric)
NASA AVIRIS, + (MODIS AIRBORNE SIMULATOR)
CASI, ULTRA-LITE

8 AIRCRAFT SENSORS

Field Measurements:

GIS information (BCMOF, BCMELP, GVWD, TFLs, etc.)
DTM (BCMELP, EMR) 5 GIS Sources
Ground calls
Ecosystem chemistry
Meteorological data

SEIDAM

TRIM #2

- ◆ At present the Surveys and Resource Mapping Branch of BCMELP has produced through contract, approximately 3300 of the estimated 7000 map sheets needed to cover the province.
- ◆ TRIM data will be available for all of the SEIDAM test sites.
- ◆ TRIM data and EMR DTMs have been combined with LANDSAT thematic mapper imagery to create baseline thematic maps at 1:250,000 scales.

SEIDAM

FOREST COVER GIS FILES

- ◆ The Inventory Branch of BCMOF maintains approximately 7,000 mapsheets at a 1:20,000 scale which describe in detail the forest, range and recreation resources.
 - Of these 7,000 mapsheets, 6,700 are within provincial forest boundaries.
 - The other 300 mapsheets are within tree farm licences, federal and provincial parks, and private lands.
- ◆ In the province there are roughly 15,000 to 20,000 openings each year which are created on terrain corresponding to roughly 2,700 mapsheets.

SEIDAM

FOREST COVER GIS FILES

- ◆ Forest land classification contains the following attributes:
 - species composition, age, height, site quality,
 - crown closure, stand structure, stand density, and history.
- ◆ The forest cover digital mapsheets also contain 16 additional overlays:
 - inventory region and compartment,
 - timber supply area (TSA) and supply block,
 - land ownership,
 - forest region and district, public sustained yield units,
 - parks, ecological reserves,
 - woodlot licences, agricultural land reserves,
 - provincial forest, range, planning cells, operability,
 - biogeoclimatic zones and recreation.

SEIDAM

FOREST COVER STANDARDS

- ◆ The standards of the estimated forest land attributes are as follows:
 - species composition is estimated to the nearest percent.
 - Currently, twenty-seven (27) commercial species
 - and eleven (11) non-commercial species are identified.
- ◆ Age is determined to the nearest year.
- ◆ Height is determined to the nearest decimeter
 - and is the mean height of the largest 100 diameters at breast height per hectare.

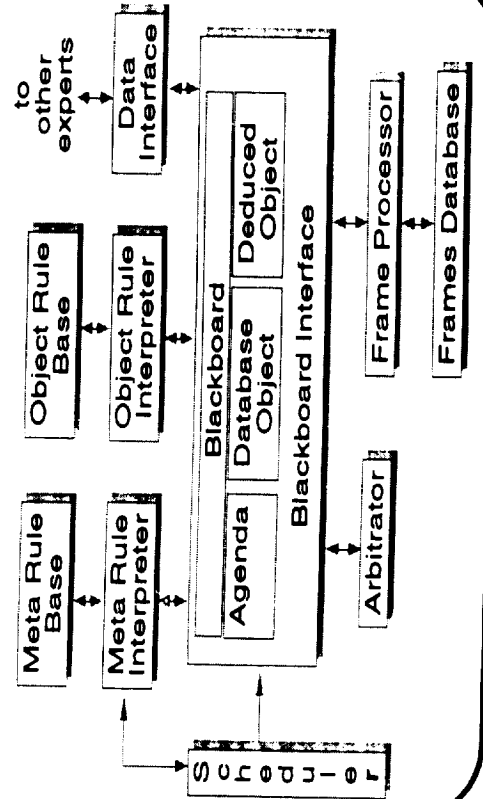
SEIDAM

FOREST COVER STANDARDS #2

- ◆ Site quality is classified as either good, medium, poor or low.
- ◆ Crown closure is the percentage of ground area that is covered by the vertically projected tree crowns.
 - It is determined to the nearest one percent, when practical; otherwise it is determined to the nearest ten percent.
- ◆ Stand structure is classified as either even aged or uneven aged; single layer or multi-layered.
- ◆ Stand density is expressed as the number of stems per hectare, above a specific diameter.

SEIDAM

RESHELL ARCHITECTURE

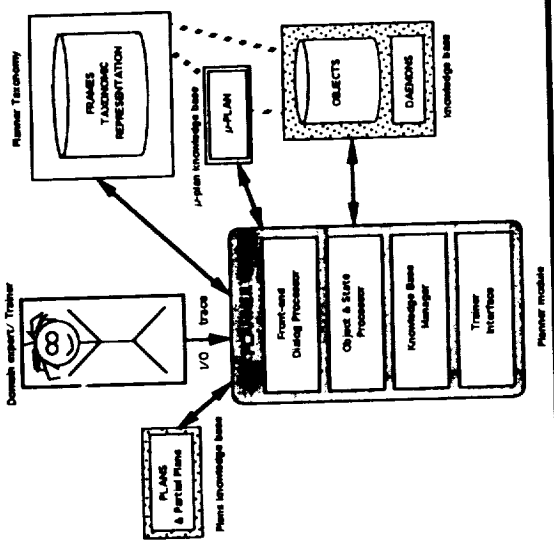


SEIDAM

THE PLANNER

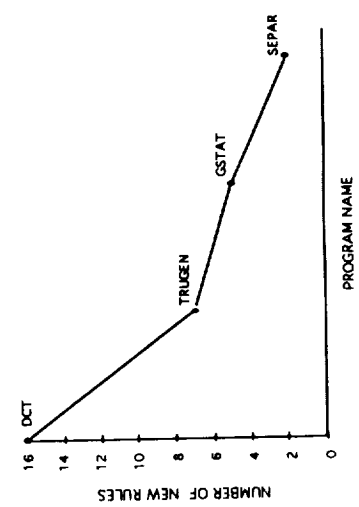
- ◆ The human expert has a goal of creating an expert system to run an existing program. He wishes to do this as simply and quickly as possible. He will use cases consisting of imagery, forest cover files, and his knowledge to create suitable state controls for the software.
- ◆ The human expert starts the planner and tells it which program is to be run. Through a dialog with the expert, and using already acquired knowledge for expert systems created earlier, the planner builds the state transitions for the given case.
- ◆ Later, rules are added to increase the expertise of the new expert system and to integrate this new system into the hierarchy of expert systems.

Learning System



LEARNING EXPERIMENT

LEARNING FOR SUPERVISED CLASSIFICATION



LEARNING REDUCES NUMBER OF NEW RULES

ML PRACTICAL ISSUES

- ◆ We have effectively used ML to speed the process of creating new expert systems at the level of controlling and running processes in a distributed computer environment.
- ◆ We want to analyze data from hundreds of images and thousands of maps (> 7,000). How do we learn from each analysis?
- ◆ If we create new knowledge, how do we ensure that this new knowledge is consistent with existing knowledge in our system?
- ◆ How do we attach ratings to the new knowledge reflecting a level of certainty in these new rules?

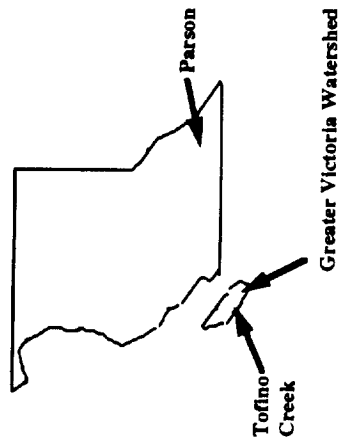
TEST SITES

- ◆ Three test sites have been selected:
- ◆ Tofino Creek (Long Beach Model Forest) - ≈ 15 km by 18 km
- ◆ Greater Victoria Watershed ≈ 15 km by 23 km
- ◆ Parson, B.C. (TFL 14) ≈ 10 km by 30 km
- ◆ In 1993 and 1994, satellite and aircraft acquisitions will take place over the test sites.

SEIDAM

LOCATION OF TEST SITES

British Columbia, Canada



SEIDAM

KNOWLEDGE ACQUISITION

- ◆ Existing TRIM and EMR DTMs have been acquired. Forest cover files and historical satellite data will be co-registered.
- ◆ In July, 1993 and 1994, satellite and aircraft acquisitions will take place over the test sites.
- ◆ Field parties will take ecological samples for foliar chemistry estimation with their locations defined by GPS ground equipment.
- ◆ Measurements will be made of chlorophyll, nitrogen, lignin, and other chemicals.
- ◆ The estimated data volume from all sources for 1993 and 1994 is 2 terabytes.

SEIDAM

KNOWLEDGE ACQUISITION #2

- ◆ Analyses will be conducted jointly with partners to establish the knowledge base for responding to queries.
- ◆ For aircraft calibration, optical measurements of lakes and homogeneous targets are made and sar corner reflectors and active radar calibrators are deployed.
- ◆ For multi-sensor registration GPS will be used.
- ◆ The Greater Victoria Watershed site has an extensive history of ecological monitoring.

SEIDAM

SEIDAM SOFTWARE STATUS

- ◆ NRCan. staff working with BCMOF created a System of Hierarchical Experts for Resource Inventories (SHERI) which updates forest GIS files with newly harvested areas derived from LANDSAT Thematic Mapper imagery.
- ◆ SHERI was installed and made operational at BCMOF in March, 1993.
- ◆ For SEIDAM, more than 65 expert systems have been created.
- ◆ Motif is used for the user interface on Unix operating system (SUN, SGI).
- ◆ Image analysis software includes CCRS LDIAS, NASA LAS, PCI, and smaller packages.

SEIDAM

SEIDAM SOFTWARE STATUS #2

- ◆ GIS software includes Pamap GIS, GRASS, and ESRI ARC/Ingres
- ◆ Data bases include Ingres and Prolog Frames data base.
- ◆ Visualization is supported by NRCan. software for generalization and AVS.
- ◆ Expert systems are built using RESHELL and the Planner.
- ◆ Expert systems can control third-party software

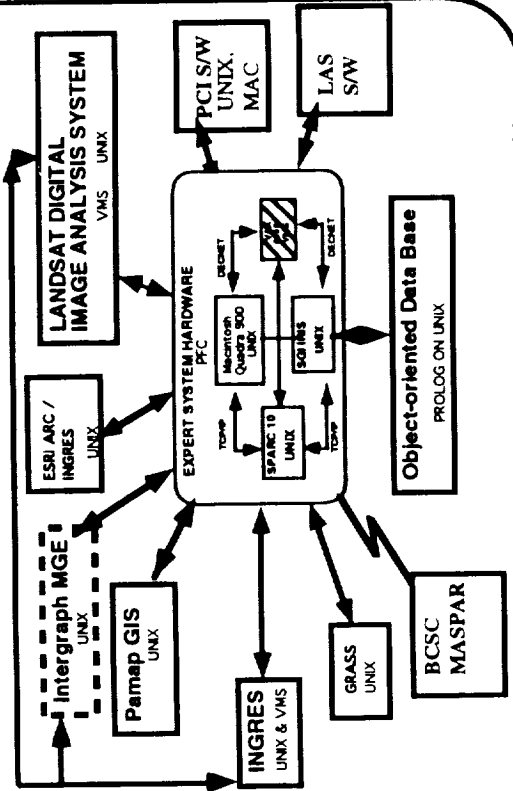
SEIDAM

FUTURE ACTIVITIES

- ◆ Research, develop and integrate new Expert Systems and implement CBR.
- ◆ Integrate forest growth and pest spread models.
- ◆ Conduct knowledge acquisition experiments for SEIDAM development and validation in 1993 and 1994.
- ◆ Demonstrate SEIDAM products for test sites, first for satellite data, then for aircraft data..
- ◆ Use Prolog Object Oriented Database.
- ◆ The Meta Data Management Subsystem is being developed now. The first release of the Meta Data Model is complete.
- ◆ Use Machine Learning to accelerate the development of new Experts and to learn from answered queries.

SEIDAM

SEIDAM - PFC COMPUTING ENVIRONMENT



SEIDAM

SEIDAM SUMMARY

- ◆ SEIDAM integrates data from multiple sources and environmental and forest models in order to respond to queries about the forests.
- ◆ SEIDAM is aimed at the future data management problems facing remote sensing.
- ◆ SEIDAM will include 150 expert systems. Machine learning will be used to expedite the creation of knowledge and new expert systems and to learn from analyses.
- ◆ SEIDAM will connect to LandData BC, the future provincial land information system.
- ◆ SEIDAM is an intelligent data management system which reduces complexity, increases adaptability, and reduces costs.

SEIDAM

OUTPUTS

- ◆ Answers to queries in the form of values, GIS files, visual images, tabular summaries, text.
- ◆ Timely resource information with accuracy on GIS labels and forest object boundaries.
- ◆ Predictive answers to queries.
- ◆ Temporal change products, canopy chemistry products.
- ◆ Approximately 150 expert systems which integrate remote sensing data, GIS files, field data through a data fusion expert to models.
- ◆ Software products and publications.
- ◆ A prototype of a system to manage large volumes of data for resource information.
- ◆ Machine learning software to create new expert systems and to learn from queries asked.

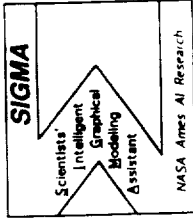
SEIDAM

SIGMA: AN INTELLIGENT MODEL-BUILDING ASSISTANT

Dr. Richard Keller
Ames Research Center

August 3, 1993

SIGMA: An Intelligent Model-building Assistant



Project Team

Richard Keller
David Thompson

Michal Rimmon
Aseem Das

Pandu Nayak

Artificial Intelligence Research Branch

Jennifer Dungan
Ecosystem Sciences

Michael Sims
Information Sciences

Chris McKay
Caitlin Griffith
Planetary Sciences

NASA Ames Research Center

Project contact: R. Keller, (415) 604-3388, keller@ptolemy.arc.nasa.gov

Keller AIRDP Workshop August 3 & 4, 1993

Outline

- I. Project Goals
- II. Model-building and difficulties with automation
- III. Knowledge for automation
- IV. Accomplishments
- V. Remaining Challenges
- VI. Testing and Technology Transfer Plans

Project Goal

1990: Develop an intelligent, domain-specific tool to assist planetary scientists in building scientific models



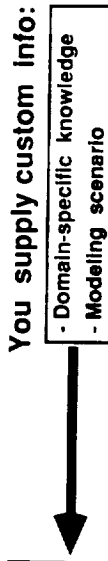
1993: Develop an intelligent, generic tool to assist scientists in a variety of domains in building scientific models

Keller AIRDP Workshop August 3 & 4, 1993

A Model-building "Shell"

We supply the shell:

- GUI
- Scientific knowledge base
 - equations
 - quantities
 - generic concepts
- Knowledge acquisition tools



Result:
Customized system

Preliminaries: What's in a model??

Equations +

Data:

$$\rho = n \sum_i f_i \frac{\mu_i}{NO}$$

$$F = G \frac{m_i n_i}{r^2}$$

$$\frac{dp}{dz} = -g r$$

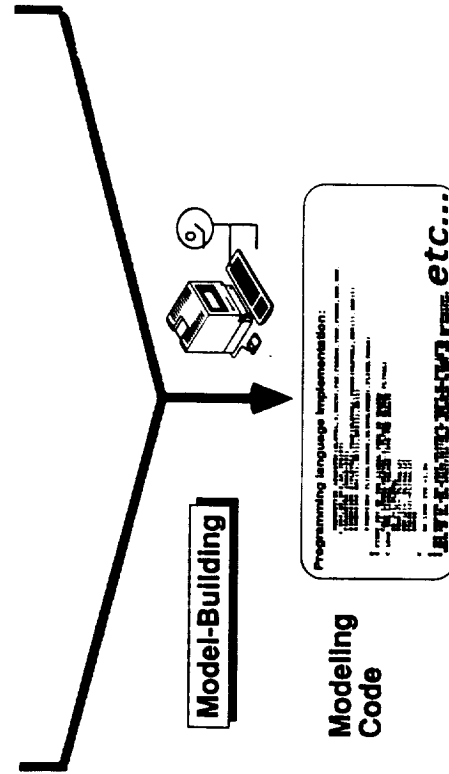
$$n = \frac{P}{kT}$$

z	n
10	3.456
20	4.553
30	15.88

$P_{tot} = 350.0$

Difficulties in Model-building

1. Do some of the symbols represent constants?
2. Do identical symbols refer to the same value?
3. What do the subscripts denote?
4. Do different symbols refer to different values?
5. What set of things does summation iterate over?
6. If differential equation can't be solved analytically, what method should be used to solve, and what are the boundary conditions?
7. Do the equations represent a valid, well-formed scientific model? Are the equations applied legitimately?
8. Which variables need to be solved for first?

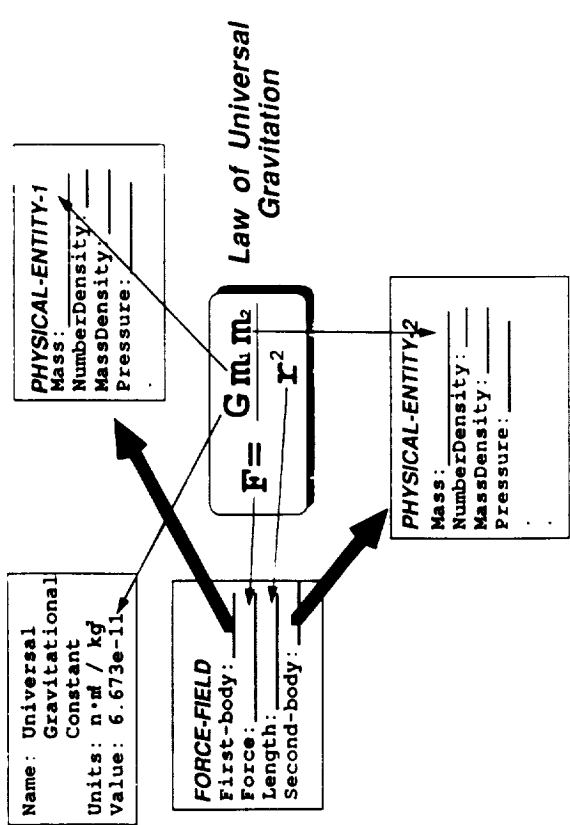


What Knowledge is required for Model-Building?

- **General knowledge about equations:**
When do equations apply in general, and what do equation symbols represent?
- **Knowledge of the modeling context:**
What is the experimental situation being modeled?
- **Articulation knowledge:** How are equations applied in the specific modeling context?

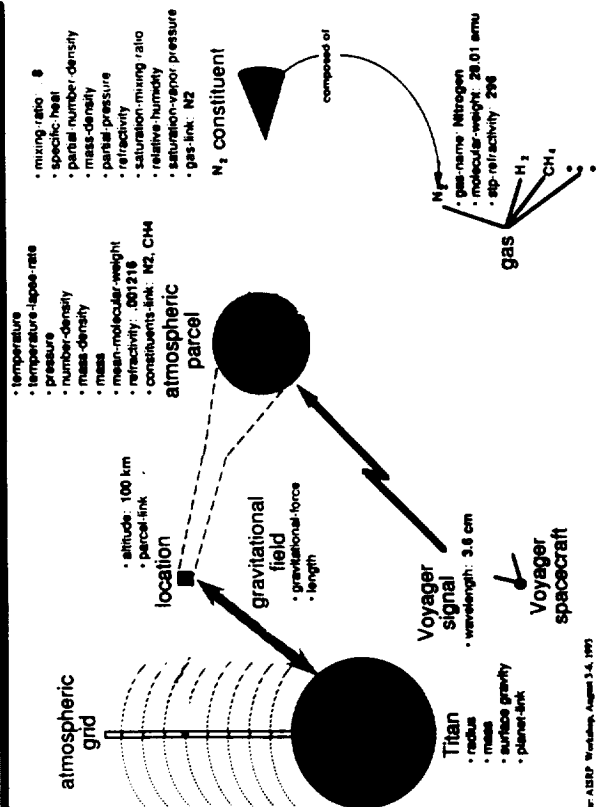
Keller, AIRSEP Workshop, August 3-4, 1993

General Equation Knowledge



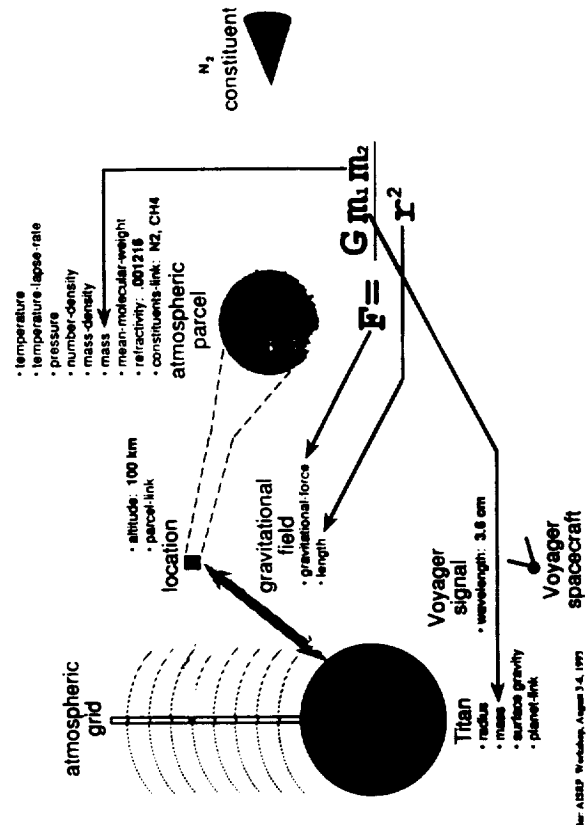
Keller, AIRSEP Workshop, August 3-4, 1993

The modeling context describes the experimental situation



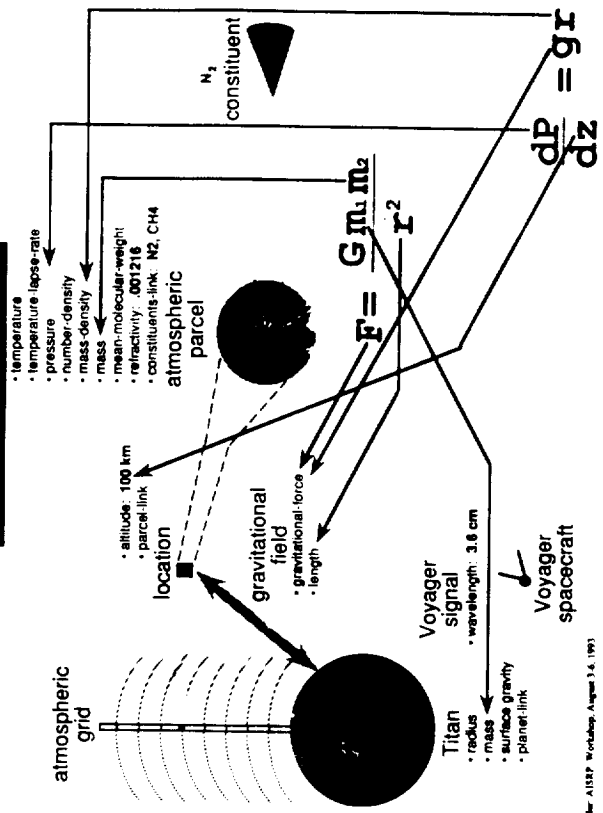
Keller, AIRSEP Workshop, August 3-4, 1993

Articulation knowledge links equations with context



Keller, AIRSEP Workshop, August 3-4, 1993

Articulation Network



Editor: AIRSP Workshop, August 3-4, 1993

How SIGMA facilitates model-building

SIGMA's interface appears similar to data flow visualization tools:

- LabView
- Khoros
- SGI Explorer

except: - SIGMA has knowledge of the scientific domain

- equation knowledge
- modeling context description

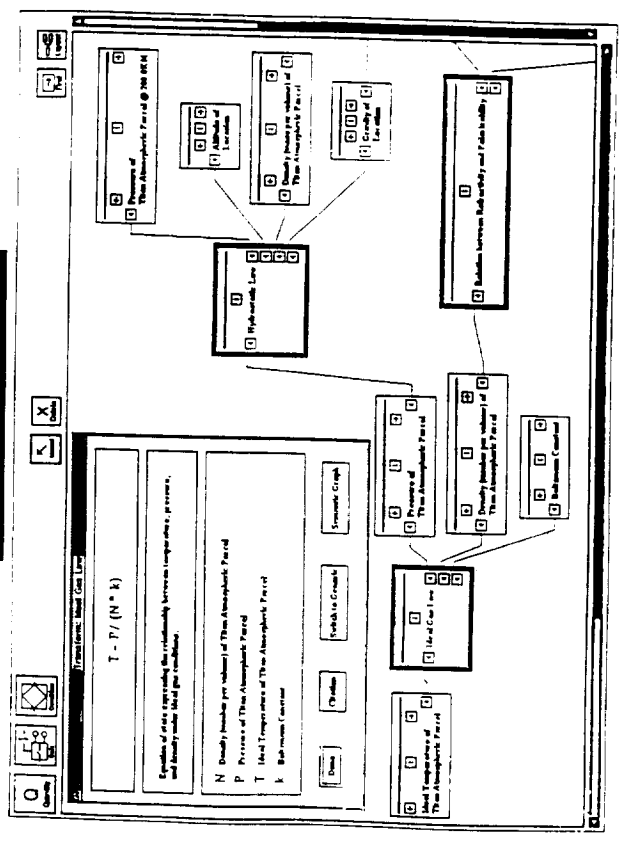
- SIGMA automatically builds an articulation network underneath the data flow using AI constraint propagation techniques

so: SIGMA users:

- receive more automated assistance
- and
- cannot build "nonsense" models

Editor: AIRSP Workshop, August 3-4, 1993

The SIGMA Interface



Editor: AIRSP Workshop, August 3-4, 1993

Accomplishments

1. The SIGMA Tool:
 - GUI for constructing, saving and restoring data flow diagrams
 - System core: constraint maintenance and propagation techniques
 - Reusable knowledge base:
 - Scientific equations and subroutines
 - Physical quantities and units
 - Domain objects and attributes
 - Scientific units conversion package
 - Literature citation facility for equations
 - Visualization facility

Accomplishments (cont'd)

2. **Model re-engineering and results replication:**
 - Lindal et al. 1983 (Voyager I data analysis)
 - Running and Coughlan 1988 (forestry data)
3. **Model-building framework (modeling as planning)**
4. **Publications**
 - Keller, Sims, Podoljak & McKay 1990
 - Keller 1991
 - Dungan & Keller 1991
 - Keller 1992a, 1992b
 - Keller & Rimmon 1992, 1993
5. **Symposia**
 - AAAI Symposium on Intelligent scientific computation
 - Panel at Automated Software Design workshop

Keller AIBSP Workshop, August 3-4, 1993

Challenges

- Knowledge acquisition and maintenance
- Scenario configuration
- Control constructs
- Coupled equations: linear, non-linear, ODEs
- Efficiency and large datasets
- Multiple, alternative models

Keller AIBSP Workshop, August 3-4, 1993

Testing

- Alpha test by collaborators currently ongoing
- New modeling tasks / participants for beta test:
 - Add soils and trace gas models to Forest-BGC
 - Add alternative haze models to Titan Greenhouse Model
 - Investigate thermal stability of Titan's atmosphere
- New modeling domain
 - Life support & exobiology: modeling bioreactor experiments involving incubation of microbial organisms

Keller AIBSP Workshop, August 3-4, 1993

Technology Transfer Plans

- University of Montana: grad student seminar
- Modelling seminar at upcoming Ecological Society of America conference
- Cassini modelers meeting
- NASA Technology Applications Team: commercialization opportunities (Research Triangle Institute)

Keller AIBSP Workshop, August 3-4, 1993

Project Bibliography

- Richard M. Keller and Michal Rimón, "A Specification Acquisition Assistant for Constructing Scientific Modeling Software", *Journal of Automated Software Engineering*, accepted for publication pending revisions, 1993.
- Richard M. Keller and Michal Rimón, "A Knowledge-based Software Development Environment for Scientific Model-building", *7th Knowledge-Based Software Engineering Conference (KBSE-92)*, Tysons Corner, VA, September 1992.
- Richard M. Keller, "Artificial Intelligence Support for Scientific Model-building", *Proc. AAAI Fall Symposium on Intelligent Scientific Computation*, Boston, MA, October 1992.
- Richard M. Keller, "Knowledge-intensive Software Design: Can too much knowledge be a burden?", *Proc. AAAI-92 Workshop on Automating Software Design*, San Jose, CA, July 1992.
- J. L. Dungan and R. Keller, "Development of an Advanced Software Tool for Ecosystem Simulation Modeling", Abstracts supplement of the *Bulletin of the Ecological Society of America*, 72(2) p.104, 1991.
- Richard M Keller, "The Scientific Modeling Assistant: An Interactive Scientific Model-Building Tool", *Proc. AAAI-91 Workshop on Automating Software Design*, Anaheim, CA, July 1991.
- R.M. Keller, M.H. Sims, E. Podolak, and C.P. McKay, "Constructing an Advanced Software Tool for Planetary Atmospheric Modeling", *Proc. I-SAIRAS'90 (International Symposium on Artificial Intelligence, Robotics and Automation in Space)*, Kobe, Japan, November 1990.

Keller, R.M. Workshop, August 1-4, 1991

MULTIVARIATE STATISTICAL ANALYSIS SOFTWARE TECHNOLOGIES FOR ASTROPHYSICAL RESEARCH INVOLVING LARGE DATABASES

Dr. S. Djorgovski
California Institute of Technology

August 3, 1993

The Motivation Behind This Effort

- The Second Palomar Sky Survey (POSS II) is now well under way; it represents a major investment of the local human, technical, and financial resources (several MB).
- In order to fully exploit the survey scientifically, the photographic plates must be digitized. Caltech and STScI producing a digital version of the survey (DPOSS).
- When complete, DPOSS will consist of ~ 37 D of pixel data an image of the entire northern sky in 3 colors, down to a limiting magnitude $D \sim 22^m$. We estimate that up to 50 billion galaxies and 500 million stars (as well as $\sim 10^6$ quasars at least $\sim 20\%$ of them at $z > 4$) will be detected in the survey.
- Processing and scientific exploitation of this vast amount of information in a systematic and uniform manner represents an unprecedented problem in astronomy (e.g., several hundred times more sources than in the entire IRAS survey). The traditional techniques are not up to the task.
- We have thus decided to explore and apply modern software technology for this purpose: machine learning, machine intelligence, expert systems, etc. An obvious source of talent and expertise was the JPL AI Group.

Cataloguing of the Northern Sky: Analysis of the Digitized Second Palomar Sky Survey

- The survey, digitization, and cataloguing
- Major science goals
- Software technology innovations
- Technology transfer to science users community

Presentation at NASA HQ, 19 April 1993
S.G. Djorgovski
MS 105-24, Caltech, Pasadena, CA 91125
Internet: george@deimos.caltech.edu
Ph. (818) 395-4415, FAX (818) 568-1517

A collaborative effort between JPL (U. Fayyad, R. Doyle, et al.) and Caltech (S.G. Djorgovski, N. Weir, et al.)

MULTIVARIATE STATISTICAL ANALYSIS SOFTWARE TECHNOLOGIES FOR ASTROPHYSICAL RESEARCH INVOLVING LARGE DATABASES

P.I. = S. GEORGE DJORGOVSKI

ASTRONOMY, MS 105-24, CALTECH, PASADENA, CA 91125
EMAIL: george@deimos.caltech.edu
Ph. (818) 395-4415 FAX (818) 568-1517

Two stages of the project:

1. Development of STATPROG, a user-friendly, science-driven package for multivariate statistical analysis of (relatively) small data sets (typically $10,000 >$ data vectors, $30 >$ dims.)
2. Development of SKICAT (nee FRITZ), an AI-based system for processing and analysis of about 3 TB of digital image information from the Second Palomar Sky Survey, and other, present and future large astronomical data sets.

SKICAT is a collaborative project between JPL (U. Fayyad, R. Doyle, et al.) and Caltech (N. Weir, S.G. Djorgovski, et al.)

- * The handout may contain additional explanatory material, not shown during the oral presentation

The Digitization of POSS-II at STScI

There are two digitization efforts under way:

1. Digitization of both the old POSS I and the new POSS II plates at USNOFS by D. Monet et al. for astronomical purposes.
2. Digitization of POSS II at STScI, by B. Laker, M. Slier, M. Vreeman, and their team

(Lower-quality glass copies of the POSS I plates have been digitized by several groups, including STScI and U. Minn.)

Features of the STScI digitization effort

- Digitization in 15 microns (1 arcsec) pixels in raster of 23.04 square, giving about 1 Gb per plate in an uncompressed form or about 3 TB of pixel data total
- Extensive quality-control tests are being performed at both STScI and Caltech
- Astrometric plate solutions will be provided for all first photographic CCD calibration efforts are now under way both STScI and Palomar
- Near completion in the 1995-1997 timeframe
- Distribution of work to the community in a compressed CD-ROM form

The Second Palomar Sky Survey

The second Palomar sky survey (POSS-II) is currently under way, using the 48-inch Oschin Schmidt telescope at Palomar Observatory. For a brief description, see Reid et al. 1991, PASP, 103, 661.

When completed, the photographic survey will consist of some 900 fields spaced at $5'$ intervals, covering the entire northern sky in three bands:

Blue-green (IIIa-J emul. + GC996 film), $\lambda_{eff} \sim 480$ nm,
Red (IIIa-F emulsion + RC610 filter), $\lambda_{eff} \sim 650$ nm,
Near IR (IV-N emulsion + RC9 filter), $\lambda_{eff} \sim 850$ nm.

At the moment, approximately 50% of the J plates, 60% of the F plates, and 15% of the N plates of adequate quality are in hand. The survey should be 90-95% complete by early 1996.

The new survey reaches $\sim 1^m$ - 1.5^m deeper than the old POSS-I, mainly due to the new fine-grain emulsion and the improved telescope optics; the gain would have been larger, but the Palomar sky is now much brighter than in 1980's. The typical limiting magnitudes are:

$D_J \sim 22.5$, $D_F \sim 20.8$, and $D_N \sim 19.5$

However, the fine-grain emulsions and the better image quality make the classification of faint objects (stars vs. galaxies) possible to a magnitude level at least 1^m - 2^m deeper than in the old survey.

Photographic copies of the survey (glass and film) are being produced at the ESO photobank in Garching, and will be distributed to subscribers at cost.

SKICAT Overview

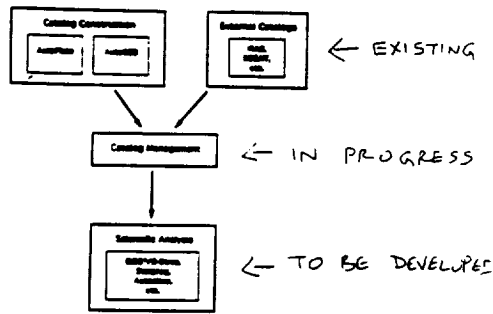
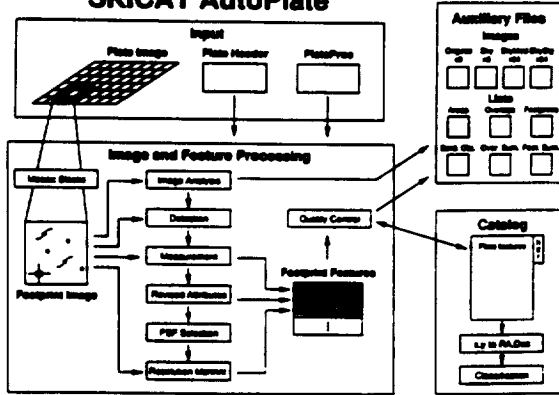


Figure 1. A schematic representation of the SKICAT system. Interfacing of the external catalogs such as IRAS, Renss, etc. (top left) with the optical data and catalogs derived from POSS-II (middle) is the subject of this proposal.

SKICAT AutoPlate



SKICAT: The programming effort so far

Language	Progr. units	Lines of code
C	400	160,000
Fortran	100	25,000
Shell scr.	100	15,000
Total:	600	200,000

Example Science Drivers

- Cataloging Efforts at Caltech**
- Using SKICAT, an AI-based system, on the cutting edge of the modern machine learning software technology
 - Processing of the incoming plate scans and CCD calibration data in a "seamless" mode, in an objective, uniform, and (semi)automatic fashion
 - The resulting Palomar Northern Sky Catalog (PNSC) will be a continuously upgraded data base, as more CCD data are added in, and accessible via computer network links - a new style for astronomical cataloging
 - We are looking into incorporating PNSC and SKICAT in the AIDS, a natural environment for them
 - The PNSC will contain up to 50 million galaxies and 500 million stars, including over a 100 thousand quasars, in 2 or 3 colors, down to the limiting survey magnitude $D_r \sim 22^m$, over the entire Northern sky
 - A fragmentary release of the PNSC may begin in the 1991-1993 timeframe, with a timely complete release within about 5 years from now
 - The PNSC will provide a scientifically rich support database for analysis of data from a wide range of space missions, as well as the planning of new ones
- Catalog of galaxies in two (three?) colors**
- Compute the galaxy-galaxy angular correlation function
 - Analyze galaxy counts and colors as a function of magnitude
 - Utilize for redshift surveys (e.g., Norris Survey)
 - Optical IDs for radio, IRAS, x-ray, etc., catalogs
 - Correlate with existing quasar catalogs
 - Catalog of compact groups of galaxies
 - Catalog of clusters of galaxies
 - Catalog of bright galaxies, including surface photometry and morphological attributes
- Catalog of point sources in two (three?) colors**
- Studies of Galactic structure
 - Search for low-mass stars
 - Search for luminous quasars
- Systematic search for low surface brightness galaxies**

SKICAT

Goals

- Automatic and optimal extraction of image information for diverse scientific programs
- Uniform object catalogs
- Interactive catalog (and ultimately image) analysis

Components

- Image processing and reduction: pixels — catalogs
- Object classification
- Catalog matching and calibration: plate catalogs + CCD catalogs — survey catalogs
- Catalog querying and analysis

Present Implementation

- Runs on Sparc II
 - FOCAS for image and initial catalog processing
 - SAS for database maintenance
- NOW
SYBASE

The Data

POSS-II

- 3 bands: IIIaJ (blue), IIIaF (red), and IV-N (near IR)
- 6.5° squared, with 5.0" spacing
- Limiting magnitudes are typically:
 $B_J \sim 22.5$, $R_F \sim 21.5$, and $I_N \sim 19.5$
- Average seeing $\sim 2''.5 - 3''.0$

CCD Calibrations

- 3 bands: Gunn g , r , and i
- Limiting magnitudes 1 - 2 mag fainter than plates
- For photometric calibration and training/test data for object classification

Scans

- Using ST-Sci PDS
- Scanning parameters:
 pixel (step) size: $15\mu\text{m}$
 aperture: $30\mu\text{m}$
- ~ 2700 images of $23,040^2$ pixels each
- 2 byte digitization - 1 Gb / plate - 3 Tb total
- Astrometric solution good to $\sim 1''$

Current Catalog Generation Scheme

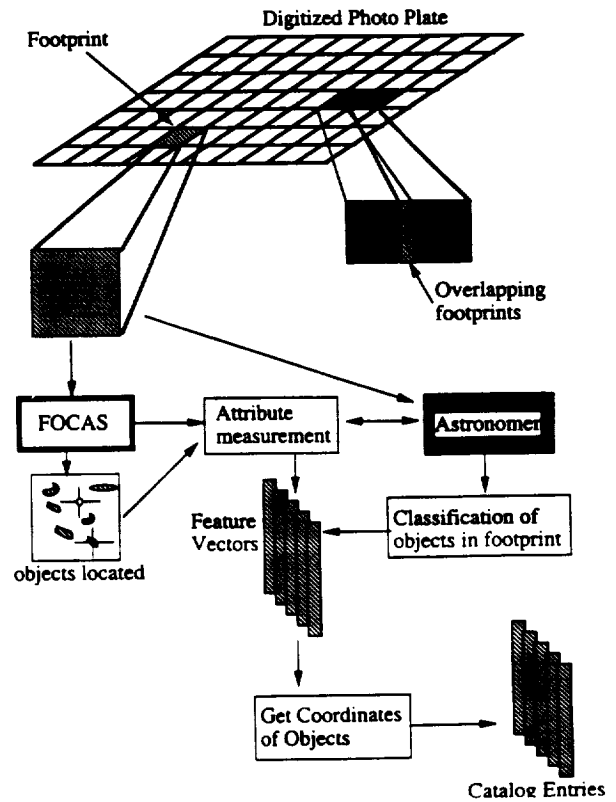


Image Processing Steps:

1. Scan (digitize) photo plate.
2. Detection: detect contiguous pixels in image that are to be grouped as one object (standard image processing).
3. Perform more accurate local sky determination for each detected object.
4. Evaluate parameters for each object independently: 40 base-level attributes.
5. Split objects that are "blended" together and re-evaluate attributes.
6. AUTOPSF: get sure-thing stars, form template.
7. Measure resolution scale and resolution fraction of each object:
 These are obtained by fitting to template of sure-thing stars.
8. Classify objects in image.

All steps are automated except for step 6.

Software used for image processing: FOCAS

When all seven steps performed, we believe from experiments so far that machine learning technique can attain more than 90% correctness in classifying all objects in the image (step 8).

So two problems to work on:

- Step 6: AUTOPSF
- final Step (8): classify objects in image.

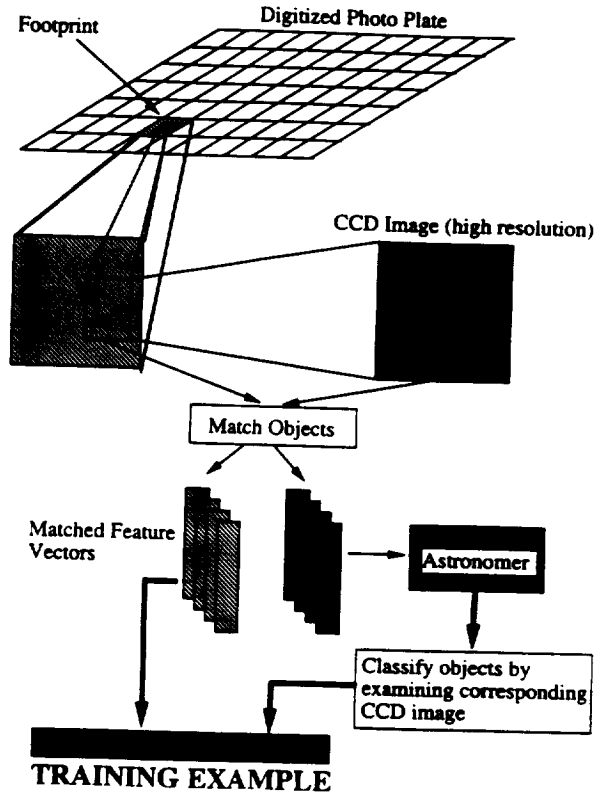
A sample of *base-level attributes* derived in Step 4 (re-evaluated in step 5):

- isophotal magnitude
- isophotal area
- core magnitude
- core luminosity
- sky brightness
- sky sigma (variance)
- image moments (8):
ir1, ir2, ir4,
r1, r2
ixx, iyy, ixy
- eccentricity (ellipticity)
- orientation
- semi-major axis
- semi-minor axis

attributes derived in step 7:

- resolution fraction
- resolution scale

Matching with Higher Resolution Sources



Have high resolution CCD frames for small areas of each plate:

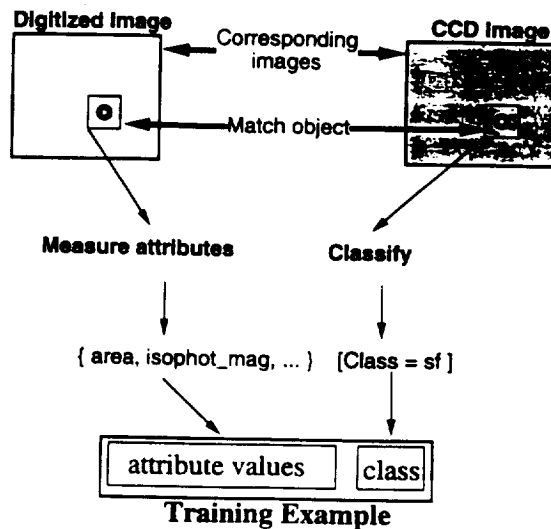
- CCD images differ only in low intensity (faint) objects
- Majority of objects on plate are faint.

Objects in CCD image can be classified by astronomer. However, many of the corresponding objects on film image are too faint to be processed by inspection.

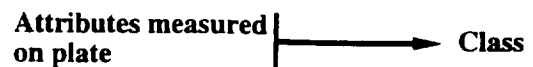
Uses of CCD frames: (other than photometry)

- Obtain classified data for training learning algorithm.
- Verify results obtained.

Classification of faint objects



Use machine learning to construct a mapping:



DECISION TREE GENERATED BY GID3

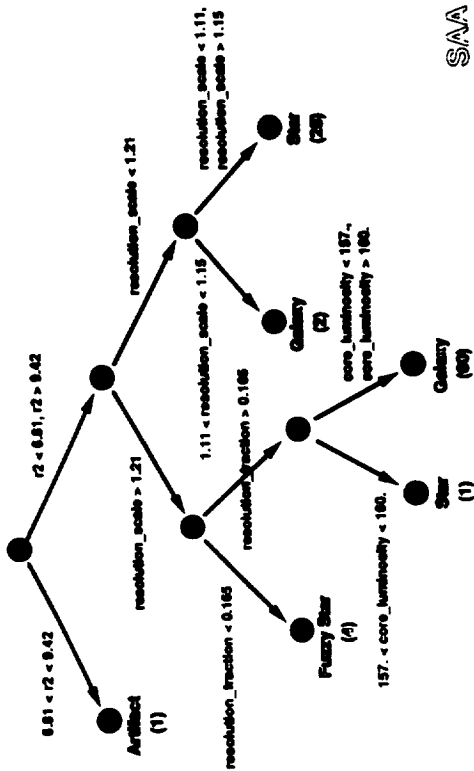
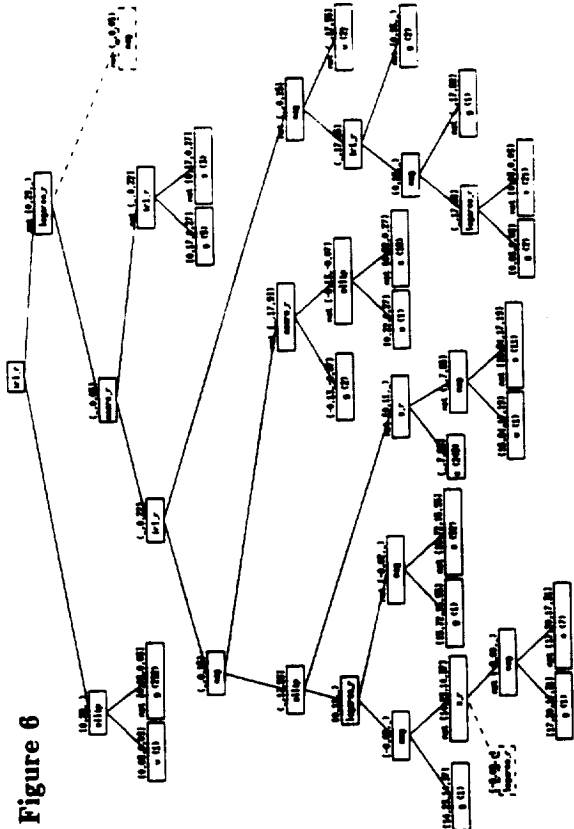


Figure 6



EXAMPLE SKICAT GENERATED DECISION TREE

Classification Results

Star Selection

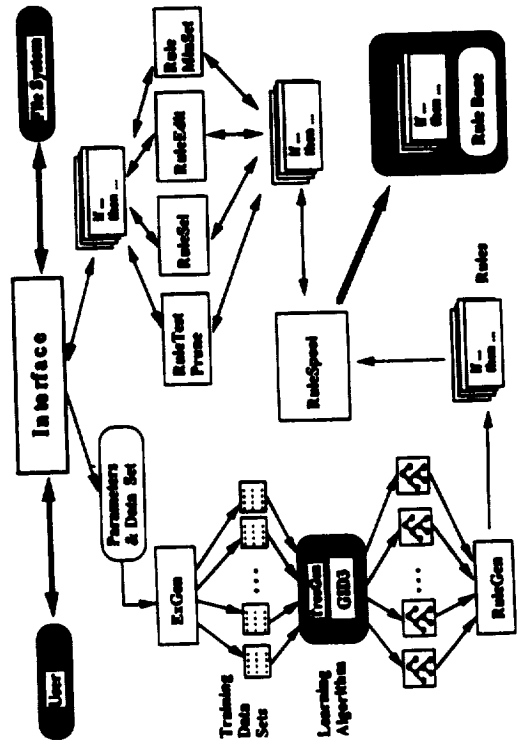
- Attributes: $m_{\text{tot}}, ell, m_{\text{core}}, \log \text{Area}, ir1,$ and $S = \text{Area} / \log[L_{\text{core}} / (9 \times isph)]$
- Error rate: $< 2\%$ for $16 < B_J < 20$
($< 5\%$ for independent COSMOS plate)

Final Classification

- Attributes:
Above plus resolution scale (α) and fraction (β)
- Error rate:

g limit	B_J limit	% error
19.5	20.2	5.1
19.9	20.6	5.4
20.2	20.9	6.5
20.5	21.2	7.0
- Results:
 - ~ 1.0 mag fainter than comparable surveys
 - effective depth $\approx 900h^{-1}$ Mpc (vs. 600)
 - ~ 1100 galaxies/deg² (vs 400)

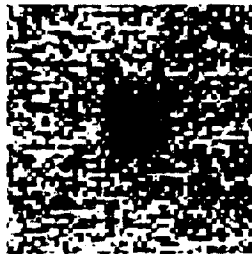
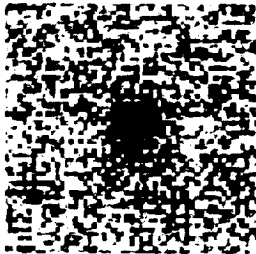
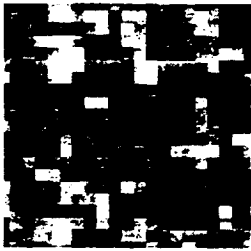
Architecture of RULER system.



POSS-II Plate J442

Star

Galaxy



Star

Galaxy

CCD Field A1632

Results of classification using all attributes:

(including resolution scale & fraction)

- used data from 3 plates in which objects were classified by astronomer after examining the corresponding CCD frames.

RULER	O-BTree	GID3*	ID3
94.1%	91.2%	90.1%	75.6%
±0.79	±1.1	±1.3	±5.1

results are for ≈ 1,100 (J380+J442) train, test within and outside plates / normalized attributes.

Neural Nets

For overall classification/star selection subproblem, used:

- traditional back propagation
- conjugate gradient optimization
- variable metric optimization

learning algorithms.

Performance comparable with decision trees, but not as stable.

Accuracy 94.2% (best) 30%(worst)
 (common range is: 76% - 84%, NC)
Avg. 81% ±5

Major Problems:

- Not interpretable
- Expensive to train:
 - slow convergence
 - how to decide number of nodes?
 - initial weight settings?

Star Selection: 97% (best),
 (common range: 78%-92%)

Why is the accurate object classification important?

An example of a scientific challenge:

We expect that the Northern Sky Catalog will contain about 500 quasars with $z > 4$. The problem is, how do we recognize them among the 200 million foreground stars?

Multicolor selection may be able to bring the number of candidate objects down to about 10,000. However, follow-up spectroscopy with a return of ~ 5% would be prohibitively expensive.

Experience by the Cambridge (APM) group suggests that most interlopers would be misclassified galaxies.

Thus, a better star-galaxy separation would make a search for high-redshift quasars much more effective!

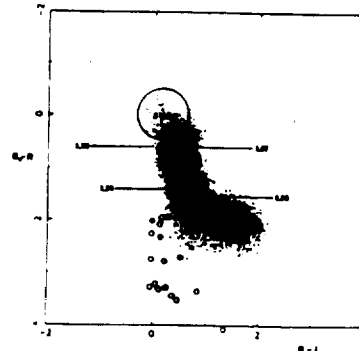


Fig. 1. A two-colour B - R, R - I plot of stellar objects in a high galactic latitude UKST standard survey field. The region occupied by the majority of quasars with $z > 4$ is shown by the large ellipse. A number of quasars

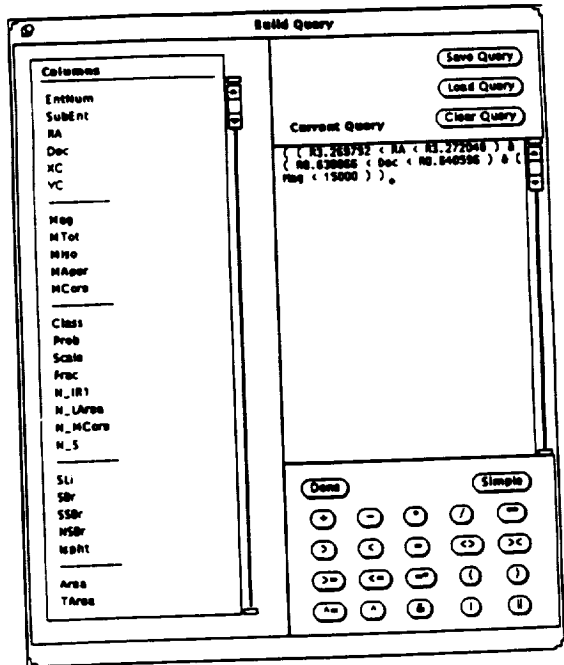


Figure 4.3: Query Table - Build Query Popop

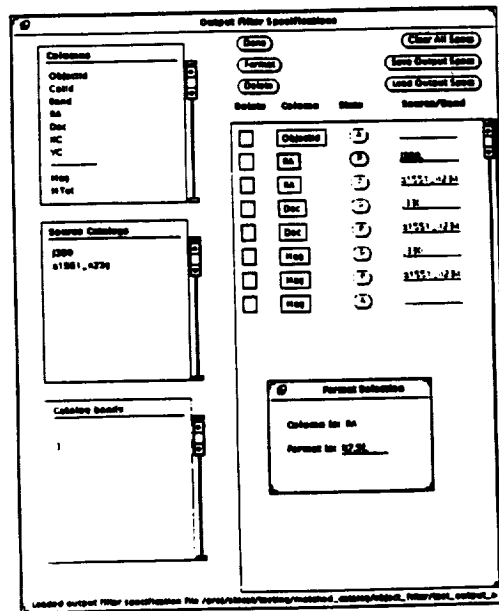


Figure 4.18: Matched Catalog Query - Output Specification Window

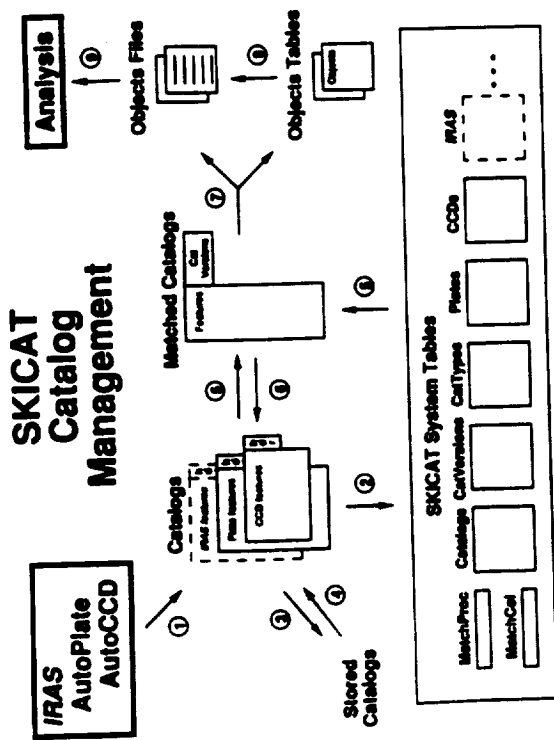


Figure 3. The catalog management module of SKICAT, which is now partly ready. The object of this proposal is to develop the appropriate interface for the external catalog of space-based data, such as IRAS or Rosat, to be automatically matched with the optical data derived from POSS-II.

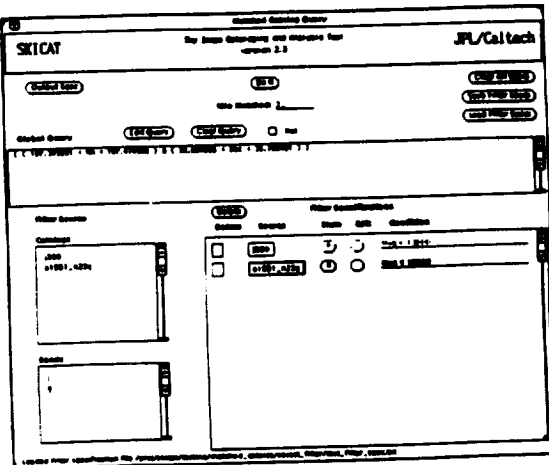
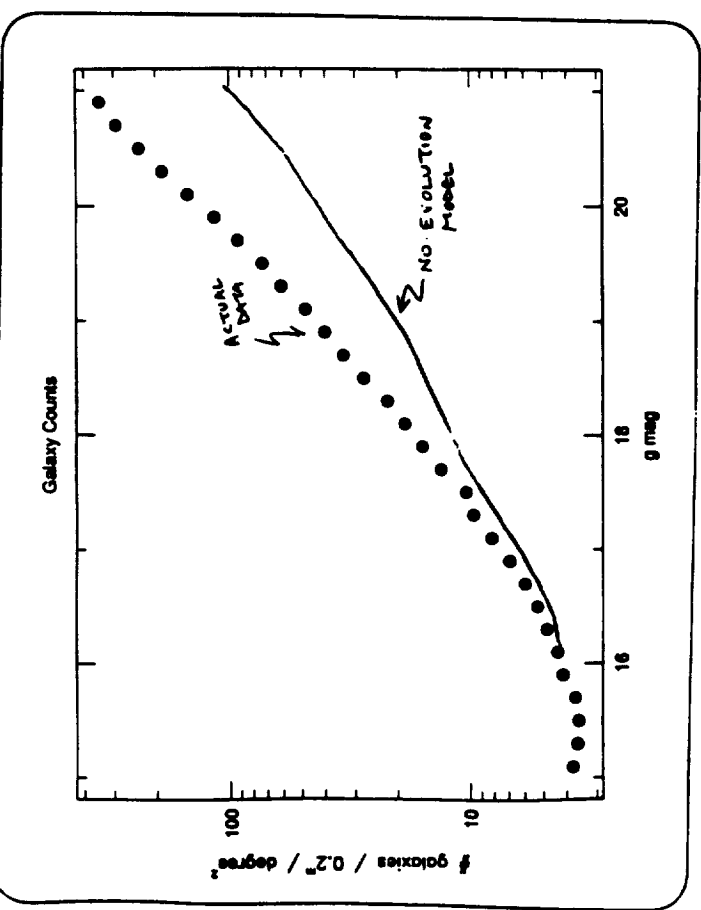
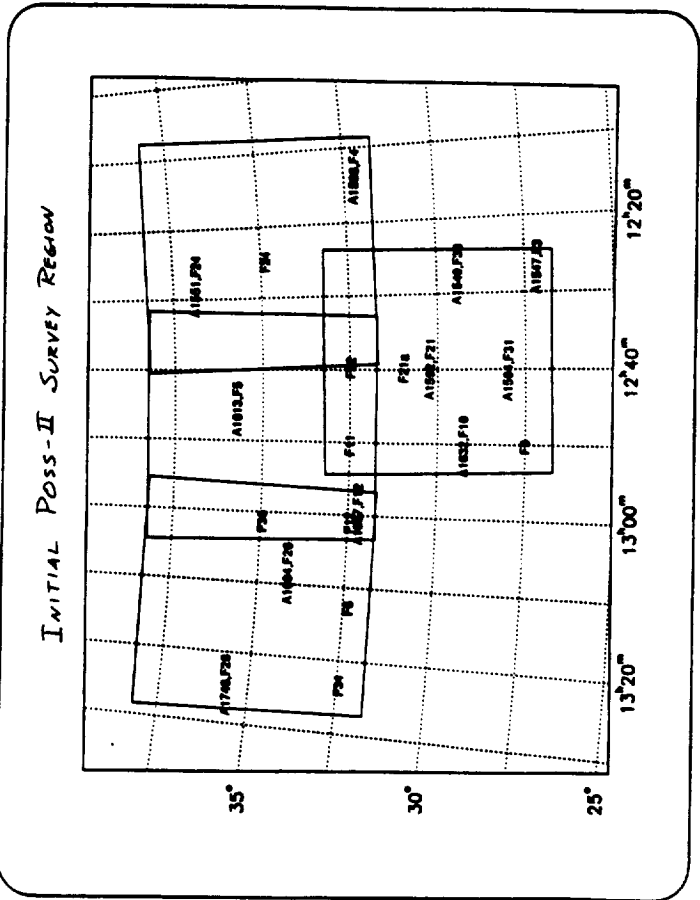
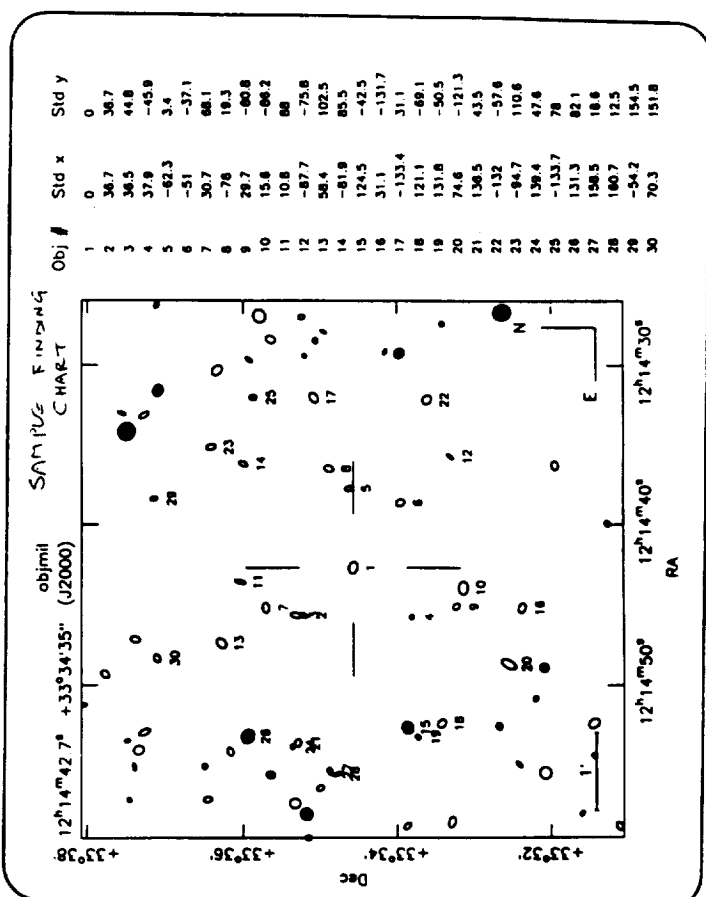
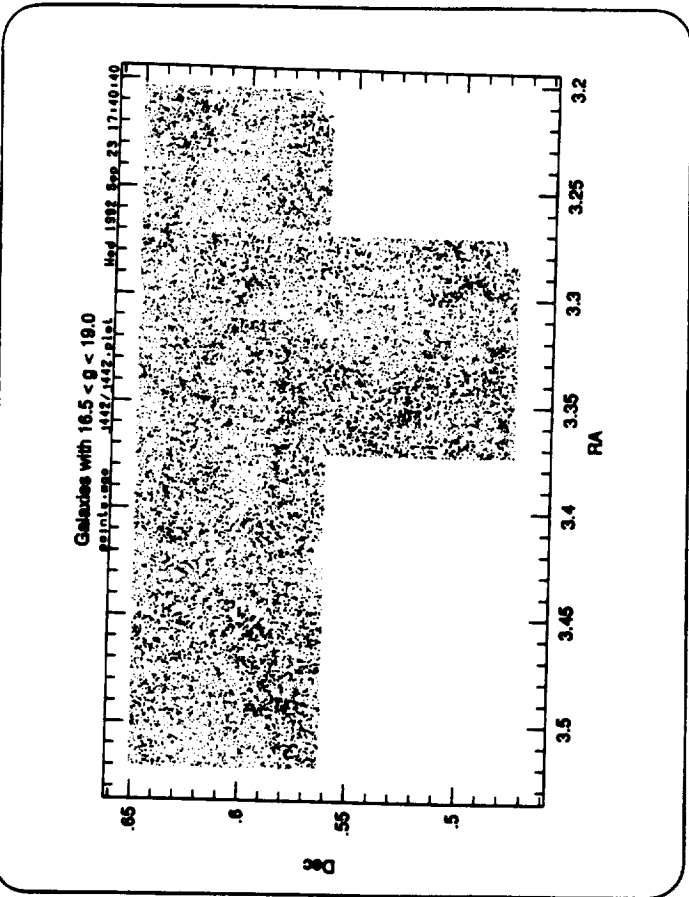
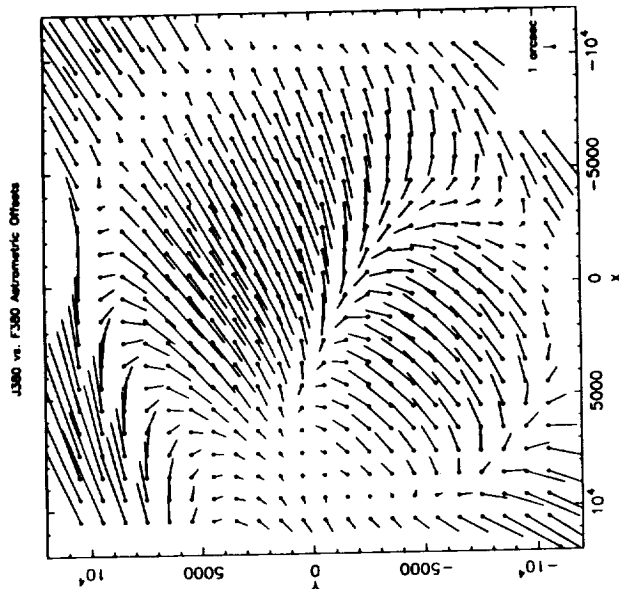


Figure 4.15: Matched Catalog Query - Main Window



ORIGINAL PAGE IS OF POOR QUALITY

QUALITY CONTROL EXAMPLE

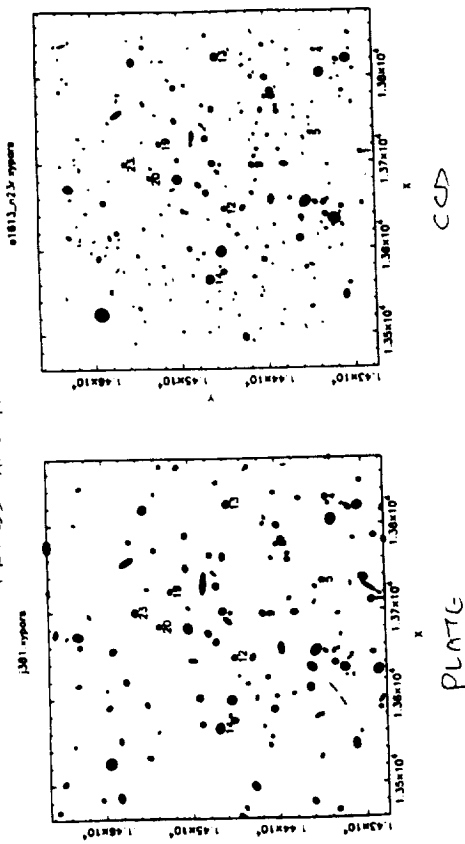


WHAT NEXT?

- We are now working on the catalog manipulation module of SKICAT, including interfaces to IRAS and other major datasets
- We are starting to explore high-level machine learning and machine discovery techniques, e.g., unsupervised classification, various Bayesian and multivariate statistical techniques (Super-STATPROG?), etc.

The initial tests using AutoClass are very encouraging - possibility of discovering new types of astronomical objects, etc.

AUTOMATIC MATCHING OF CCD CALIBRATION FRAMES AND PLATE SCANS



THE STATUS OF THE PROJECT:

- STATPROG in a regular scientific applications use, produces good science, new results
 - Major parts of SKICAT completed and tested, documentation being completed and improved
 - Some general utilities developed:
 - Data visualization: PONGO, a general, public-domain graphics package, a blend of the two most popular (in astronomy) graphics packages, PGplot and Mongo.
 - X-windows based database query user interface (DB engine driver)
 - General catalog manipulation tools for SAS and SYBASE DB environments
 - Finding charts graphing tools etc. etc.
 - Initial scientific verification and testing of SKICAT started:
 - Plate scans quality control feedback to STScI
 - Faint galaxy counts and galaxy evolution at moderate redshifts (Weir, SGD)
 - Large-scale structure and galaxy correlations analysis (Brainerd, Weir, SGD)
 - Searches for high-redshift quasars (Smith, Weir, SGD)
- ... and much more to come!

The Benefits of This Effort: Present and Future

We have successfully combined the technical expertise of the JPL AI group with the scientific applications expertise in the Caltech Astronomy Department.

The result is an unprecedented astronomical software system, SKICAT, which draws on the latest in machine learning and expert systems software technology, and which will:

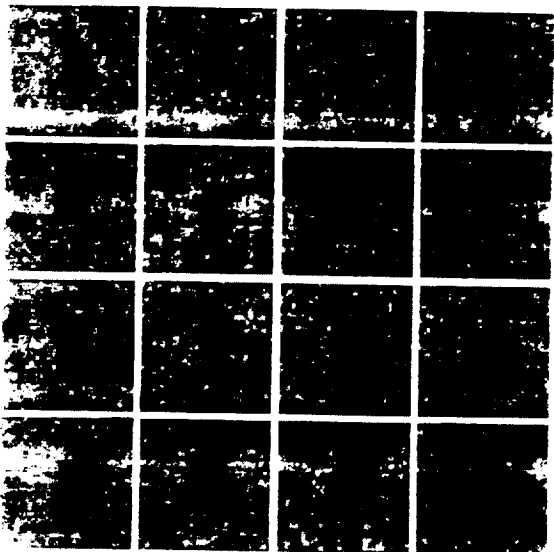
- In the short term lead to an efficient scientific exploitation of the Digitized POSS-II, and the generation of PINS
- Provide a set of powerful, modern software tools for scientific exploitation of the PINS and a plethora of other major astronomical catalogs (e.g., IRAS, ICRF, etc.)
- Possibly become a standard package to deal with large astronomical imagery data sets, both from the space and ground-based observations
- In the intermediate and long term, provide a ready infrastructure technology support for the forthcoming digital sky surveys of the future, e.g., the Two-Micron Sky Survey, surveys for the Earth-crossing asteroids, the DSS, etc.

We have thus bridged an important gap between computer-science experts and scientific applications end-users. We believe that this is a model effort in the effective technology transfer for astronomy and space science.

Long-Term Plans: Think Big

- Our long-term goal is to explore and develop methods of astrophysical research involving very large data bases, and to open up the use of the rapidly developing expert systems and related software technology in astronomy and space science.
- We are envisioning a next-generation astronomical software utility package, an order of magnitude more complex and more powerful, whose development will also teach us about the prospects and limitations of AI tools in astronomical applications.
- We are really proposing to generalize the concept of an astronomer's desktop, from that of a fixed, or semi-fixed, uplink link of objects, to a dynamical, evolving entity, which can constantly improve as better information becomes available.
- This process of an evolving catalog actually leads to a better understanding of the input data and their limits, and, as by-product, provides constant and uniform quality control, far more than our immediate applications.
- The essence of this effort is to introduce expert systems and machine learning techniques in the analysis of modern astronomical data in a major way.
- We are our effort as a development of software tools for the astronomy of the turn of the century, and beyond.

Examples from 4 Classes Discovered by AutoClass



Bibliography of the papers based wholly or in part on the work done in this project, as of July 1993:

1. Papers involving STATPROG library and related work:

- "The Manifold of Globular Clusters", Djorgovski, S. 1991, in James, K. (ed.), *Formation and Evolution of Star Clusters*, A.S.P. Conf. Ser. 13, 112.
- "Systematic Differences Between the Field and Cluster Ellipticals", de Carvalho, R., and Djorgovski, S. 1992, *Astrophys. J. Letters* 389, L49.
- "On the Formation of Globular Clusters in Elliptical Galaxies", Djorgovski, S., and Santiago, B.X. 1992, *Astrophys. J. Letters* 391, L85.
- "Dynamical Evolution Effects on the Hot Stellar Populations in Globular Clusters", Djorgovski, S., and Piotto, G. 1992, *Astrophys. J.* 104, 2112.
- "Galaxy Manifolds and Galaxy Formation", Djorgovski, S. 1992, in G. Longo et al. (eds.), *Morphological and Physical Classification of Galaxies*, p. 337. Dordrecht: Kluwer.
- "Systematics of Galaxy Properties: Clues About Their Formation", Djorgovski, S. 1992, in R. de Carvalho (ed.), *Cosmology and Large-Scale Structure in the Universe*, A.S.P. Conf. Ser. 24, 19.
- "Properties of Dwarf Spheroidals", Djorgovski, S., and de Carvalho, R. 1992, in G. Longo et al. (eds.), *Morphological and Physical Classification of Galaxies*, p. 379. Dordrecht: Kluwer.
- "Properties of Brightest Cluster Members", Djorgovski, S., de Carvalho, R., Shlosman, I., Schombert, J. 1992, in G. Longo et al. (eds.), *Morphological and Physical Classification of Galaxies*, p. 427. Dordrecht: Kluwer.
- "Systematic Differences Between the Field and Cluster Ellipticals", de Carvalho, R., and Djorgovski, S. 1992, in B. Barbuy and A. Renzini (eds.), *The Stellar Populations of Galaxies*, Proceedings of the IAU Symp. #149, p. 400. Dordrecht: Kluwer.
- "Surface Photometry of Southern Ellipticals and Fundamental Plane Solutions for the Fornax Cluster Galaxies", Penereiro, J., Djorgovski, S., de Carvalho, R., Gorjian, V., Thompson, D. 1992, in R. de Carvalho (ed.), *Cosmology and Large-Scale Structure in the Universe*, A.S.P. Conf. Ser. 24, 123.
- "Surface Photometry and Fundamental Plane Solutions for Elliptical Galaxies in the Virgo and Coma Clusters", Gorjian, V., Djorgovski, S., de Carvalho, R., Penereiro, J., Weir, N. 1992, in R. de Carvalho (ed.), *Cosmology and Large-Scale Structure in the Universe*, A.S.P. Conf. Ser. 24, 129.
- "The Manifold of Low Surface Brightness Dwarf Galaxies", de Carvalho, R., and Djorgovski, S. 1992, in R. de Carvalho (ed.), *Cosmology and Large-Scale Structure in the Universe*, A.S.P. Conf. Ser. 24, 135.
- "Multivariate Analysis of Hickson's Compact Galaxy Groups", Djorgovski, S., Weir, N., and de Carvalho, R. 1992, in R. de Carvalho (ed.), *Cosmology and Large-Scale Structure in the Universe*, A.S.P. Conf. Ser. 24, 141.
- "On the Effects of Cluster Density and Concentration the Horizontal Branch Morphology: The Origin of the Blue Tails", Fusi Pecci, F., Ferraro, F., Bellazzini, M., Djorgovski, S., Piotto, G., and Buonanno, R. 1993, *Astrophys. J.* 108, 1145.
- "What Determines the Stellar Mass Functions in Globular Clusters?", Djorgovski, S., Piotto, G., and Capaccioli, M. 1993, *Astrophys. J.* 108, 2148.
- "Multivariate Analysis of Globular Cluster Systems in Early-Type Galaxies", Santiago, B.X., and Djorgovski, S. 1993, *M.N.R.A.S.* 261, 753.

Annual Workshop on Space Operations, Applications, & Research (SOAR-92). NASA CP-3187, p. 340.

- SKICAT: A Machine Learning System for Automated Cataloging of Large Scale Sky Surveys". Fayyad, U.M., Weir, N., and Djorgovski, S. 1993, in *Proceedings of the Tenth International Conference on Machine Learning*, p. 112. San Mateo, CA: Morgan Kaufmann Publ.
 - Automated Analysis of a Large-Scale Sky Survey: The SKICAT System". Fayyad, U.M., Weir, N., and Djorgovski, S. 1993, in G. Pisotaky-Shapiro (ed.), *Proceedings of AAAI-93 Workshop on Knowledge Discovery in Databases*, p. 1. Washington: AAAI Press.
 - SKICAT: A Cataloging and Analysis Tool for Wide Field Imaging Surveys" Weir, N., Fayyad, U., Djorgovski, S., Roden, J., and Rouquette, N. 1993, in R. Hanisch, R. Brinenden, and J. Barnes (eds.), *Astronomical Data Analysis Software and Systems II, A.S.P. Conf. Ser.*, in press.
 - Automated Cataloging and Analysis of Sky Survey Image Databases: the SKICAT System". Fayyad, U., Weir, N., & Djorgovski, S. 1993, in *Proc. of the 2nd International Conference on Information and Knowledge Management (CIKM-93)*, Washington: ISCA/ACM, in press.
 - Digitized POSS-II: Galaxy Number Counts in Two Colors Over a Multi-Plate Region". Weir, N., Djorgovski, S., and Fayyad, U. 1993, *Bull. Am. Astron. Soc.* 24, 890.
 - Cataloging of the Northern Sky from the POSS-II using a Next-Generation Software Technology". Djorgovski, S., Weir, N., and Fayyad, U. 1993, IAU Commission 9 Working Group on Wide-Field Imaging Newsletter #4, 14.
 - Analysis of the Palomar-STScI Digital Sky Survey: an Overview of the SKICAT System". Weir, N., Djorgovski, S., Fayyad, U., and Roden, J. 1993, IAU Commission 9 Working Group on Wide-Field Imaging Newsletter #4, 15.
 - The Initial Results on Galaxy Counts and Searches for $z > 4$ Quasars using the Palomar-STScI Digital Sky Survey". Weir, N., Djorgovski, S., and Fayyad, U. 1993, IAU Commission 9 Working Group on Wide-Field Imaging Newsletter #4, 54.
- (The last three references will appear in a more extended form as contributed papers in the proceedings of IAU Symposium 161.)

ORIGINAL MADE IS
OF POOR QUALITY

THE GRID ANALYSIS AND DISPLAY SYSTEM (GrADS): A PRACTICAL TOOL FOR EARTH SCIENCE VISUALIZATIONS

**Dr. James L. Kinter III
University of Maryland**

August 4, 1993

**THE GRID ANALYSIS AND DISPLAY SYSTEM
(GrADS):**

**A PRACTICAL TOOL FOR
EARTH SCIENCE VISUALIZATION**

**BRIAN E. DOTY
JAMES L. KINTER III**

**CENTER FOR OCEAN-LAND-ATMOSPHERE INTERACTIONS
DEPARTMENT OF METEOROLOGY
UNIVERSITY OF MARYLAND
COLLEGE PARK, MD 20742**

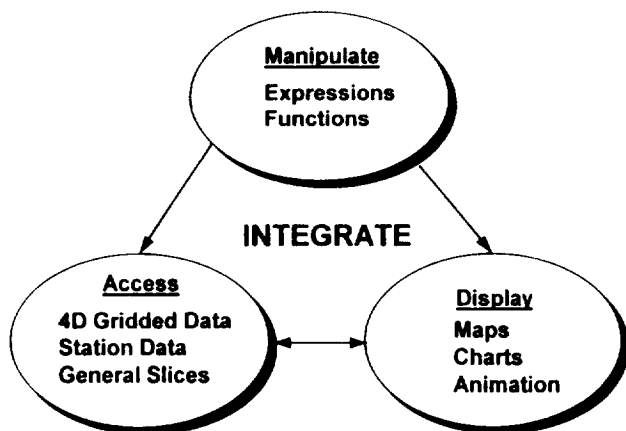
(301) 405-5394 (5395)

bdoty@cola.umd.edu

zkinter@cola.umd.edu

**SUPPORTED BY APPLIED INFORMATION SYSTEMS RESEARCH PROGRAM
NASA HEADQUARTERS**

GrADS Design Goals



GrADS Design Goals

INTERACTIVE

**Full Control of Data and Display
Sub-Second Response Time
Customizability - Scripting Language**

EASY TO LEARN AND USE

HARDCOPY VIA VECTOR GRAPHICS

Interpreted Command Line Scripting Language

Language Design:

- Programmability: as simple as possible
- Form GrADS commands via string manipulations and pass back to program for execution
- Return command results as script variables

Language Elements:

- Variables of type "character"
- Arithmetic and logical operators
- Built-in and user-specified functions
- Flow control: loops, if/then/else
- Fully recursive

Sample Usage:

- Automate commonly used command sequences
- Perform complex calculations
- Create new GrADS data files from results of GrADS calculations
- Interact with the graphics screen

GrADS DATA SETS

- **FORMATS:** GrADS INTERNAL -OR- GRIB -OR- DRS
 - Binary
 - Optimized for I/O Performance
 - Support for arbitrary length file/record headers
- **CREATE**
 - Fortran OR C
 - Standard I/O Statement
- **MODIFY**
 - Fortran or C or UNIX file commands
 - Update in place
 - Extend
- **USE IN OTHER APPLICATIONS**
 - Fortran or C
- **PORTABILITY**
 - All UNIX Computers (E. G. NFS)
 - DOS-based personal computers
- **OTHER FORMATS CAN BE SUPPORTED**
(packed binary, ASCII, net CDF, etc.)

TECHNOLOGY TRANSFER - The GrADS Experience

- There are hundreds of GrADS users at over 30 institutions
- They use GrADS for a wide variety of scientific investigations
- They haven't complained much

WHY?

- GrADS is easy to use and learn by design
- GrADS allows scientists to access/manipulate/see their data without requiring additional programming
- It is easy for scientists to get their data into and out of GrADS
- We have been willing to work with user groups to add the functionality they request; we have paid attention to user feedback and eliminated their "show stoppers"
- We have packaged the software for distribution to make installation simple for a scientist ==> no system administrator assistance is required
- GrADS is platform independent - it runs on Unix workstations and PCs so scientists can do their work wherever it suits them to do it
- GrADS is distributed free of charge

CURRENT GrADS USAGE

Research:

- Model output analysis
 - Global atmospheric general circulation models
 - Global ocean models
 - Tropical models
 - Coupled ocean-atmosphere models
- Observational data analysis
 - Station data (African rainfall, Asian soil moisture, etc.)
 - Gridded objective analyses

Education:

- Interactive classroom use
- Student research projects
- Student self-education

Forecasting:

- Real time observational data analysis

Public information:

- Daily weather forecasts
- Maryland state ozone maps
- Seminars with interactive displays

ACKNOWLEDGEMENT

Data and script graciously provided by

Dr. Richard Reynolds

NOAA/NMC/Climate Analysis Center

Data Set: Weekly sea surface temperature (SST) analyses

Script: Interactive selection of domain, time period;
animate results

ACKNOWLEDGEMENT

Data and script graciously provided by

Dr. Steven Lord

NOAA/NMC/Development Division

Data Set: 12 hourly AVN model forecasts

Script: Interactive selection of domain, time period,
variables; animate hurricane tracks

ACKNOWLEDGEMENT

Data and script graciously provided by

Roy Jenne

NCAR/SCD/Data Support Section

Data Set: CD-ROM with 12 hourly NMC analyses

Script: Data Graphical User Interface prototype

ACKNOWLEDGEMENT

Data and script graciously provided by

Dr. James Carton

University of Maryland
Department of Meteorology

Data Set: Monthly Atlantic Ocean temp. OI analyses

Script: Interactive selection of vertical level; displays
vertical temperature sections in a small region

ACKNOWLEDGEMENT

Data and script graciously provided by

Mr. John Garthner

Fleet Numerical Oceanography Center

and

Lt. Michael Fiorino

Naval Oceanography Reserve Activity 0966

Data Set: Daily Navy Operational Global Atmospheric Prediction System (NOGAPS) analyses and forecasts with surface observations

Script: GrADS interface to NODDS - Navy Oceanographic Data Distribution System

ACKNOWLEDGEMENT

Data and script graciously provided by

Dr. Paul Dirmeyer

University of Maryland
Department of Meteorology

Data Set: Hourly Surface Airways station observations

Script: Interactive selection of domain, time, variable; objectively analyzes station data and displays maps or station time series

ACKNOWLEDGEMENT

Figure graciously provided by

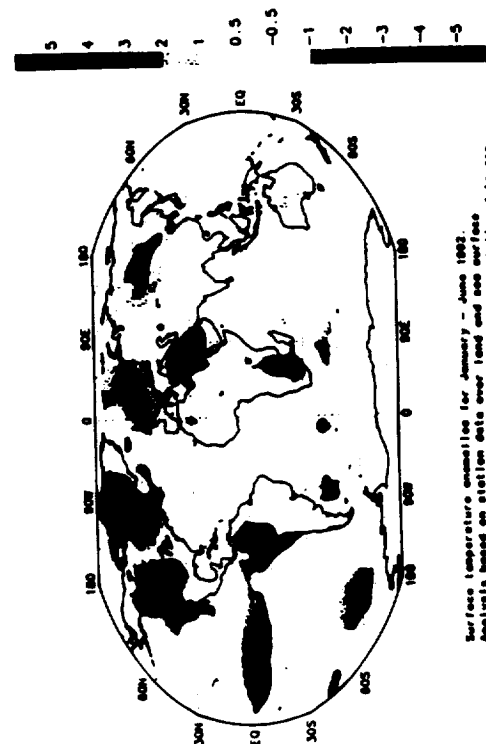
Chester Ropelewski

and

Michael Halpert

NOAA/NMC/Climate Analysis Center

Figure: Excerpt from Fourth Annual Climate Assessment: Surface temperature anomalies for January - June 1992.



Major New Features

- GrADS Scripting Language (GSL) Enhancements
 - Global Variables
 - Compound Variables; Arrays
- Batch Mode
- Shell Escape
- Support for GRIB Data Format
- Multiple File Time Series
- Additional Functions
 - Finite Differencing Functions
 - skip, const
- User Defined Functions
- Additional Graphics
 - Station Models
 - Label Formats (Axis and Contour Labels)
 - Enhancements to Colorized Vectors, Streamlines
- General Coordinate Transformations
- Data Editing

SUMMARY

CUSTOMIZABILITY

- Scripting language added during past year permits users to *customize* virtually all aspects of GrADS functionality

INTERACTIVE USAGE

- Scripting language allows interaction with display for interactive scientific investigation using one's own data

USER COMMUNITY

- GrADS is being used by a growing community of users nationally and internationally
- User feedback has been extremely helpful in prioritizing development tasks and initiating new ideas

PLANS

- Other data format standards, GUIs for specific data sets

FOR MORE INFORMATION

Contact:

BRIAN DOTY ... doty@cola.umd.edu

A DISTRIBUTED ANALYSIS AND VISUALIZATION SYSTEM FOR MODEL AND OBSERVATIONAL DATA

Dr. Robert Wilhelmson
National Center for Supercomputing
University of Illinois at Urbana-Champaign

August 4, 1993

● A Distributed Analysis and Visualization System for Model and Observational Data ●

Robert Wilhelmson

Crystal Shaw

David Wojtowicz

National Center for Supercomputing Applications (NCSA)

University of Illinois at Urbana-Champaign

John Hagedorn

NASA - Goddard Space Flight Center



A Distributed Analysis and Visualization System for Model and Observational Data

The project is part of a larger effort called PATHFINDER (Probing ATmospHeric Flows in an INteractive and Distributed EnviRonment) at NCSA. It is designed to bring to the researcher a flexible, modular, collaborative, and distributed environment for use in studying atmospheric and fluid flows, and which can be tailored for specific scientific research and weather forecasting needs.

The topics covered here include :

- PATHFINDER Objectives
- What We Have Accomplished
- Deliverables and Time Tables
- The Future
- PATHFINDER Examples

* NCSA Mosaic is used as the software for this talk.

PATHFINDER Objectives:

- To utilize existing diagnostic and analysis software capabilities such as those found in GEMPAK (the GEneral Meteorological PACkage built at Goddard Space Flight Center)
- To access a variety of analysis and display capabilities including three dimensional rendering, animation, and collaborative tools for interacting with remote users on different workstations across the national network
- To couple multiple and heterogeneous computers (e.g. SGI VGX, Cray Y-MP, CM-2, CM-5)
- To exploit the use of off-the-shelf software, both commercial and public domain: IRIS Explorer, Inventor, AVS, HDF, DTM
- To manage large amounts (gigabytes and beyond) of data generated by satellites, observational field programs, and model simulations
- To incorporate video recording and playback, high definition monitors, and a virtual reality viewer
- To process multiple data streams (both model and observational) in creating a visualization

Software Environment



Explorer

- is the primary development and execution environment
- is highly interactive
- supports distributed computing
- is easily extensible
- provides user interface tools
- has a rich collection of existing modules
- NCSA is working closely with Explorer development team



Inventor

- is an object-oriented 3-D toolkit
- provides a library of objects that you can use, modify and subclass
- simplifies application development
- facilitates moving data between applications with 3-D Interchange File Format



- is used to convert a subset of PATHFINDER modules on Convex and SGI machines
- is highly interactive
- supports distributed computing
- is easily extensible
- provides user interface tools
- has a rich collection of existing modules



HDF

- is supported on a lot of platforms
- is a multi-object file format
- has Fortran and C calling interfaces for storing and retrieving 8- and 24-bit raster images, palettes, scientific data and annotations
- is the standard archive format for EOSDIS



DTM

- uses TCP/IP for transmission of data
- is built on top of Berkeley sockets interface
- has Fortran and C calling interface
- is easy to use

What We Have Accomplished


- GEMVIS
- Animation Tool
- Metadata
- Data Acquisition
- Visual Idioms
- Laser Disk Recording
- Prototypes
- Explorer Alpha and Beta Testing

GEMVIS

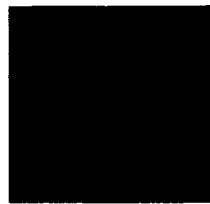


- A subset of PATHFINDER
- A set of capabilities targeted to the GEMPAK user community :
 - GEMPAK grid diagnostics & file access :
 - the full range of GEMPAK functions and operators supported
 - a new operator implemented for 3D vector data
 - all GEMPAK grid data formats supported

- all vertical coordinate systems handled
- relevant modules:
 - GemGetGrid , GemGridFunction
- map projections & topography :
 - a wide variety of map projections
 - maps registered with data set visualizations
 - topography displayed in appropriate projections
 - relevant modules:
 - GemDraw Map, GemTopo, GemGetGrid, and GemGridMapping

- Implemented in both IRIS Explorer and AVS to reach widest audience 
- Packaged for easy use by current GEMPAK users
- To be released through Unidata & COSMIC

Animation Tool



(click to enlarge)

- Is an Inventor based interactive tool for composition of complex 3d animations
- Uses keyframes to specify the animation of various scene parameters including:
 - Camera position
 - Object Movement
 - Object Material
 - Lighting/Environment
- Can be linked to Explorer to animate Explorer output over:
 - Time
 - Analysis parameters such as isosurface threshold
- Can produce raw image files for each frame or SGI/MPEG movies.
- Is based on Explorer's Render module making it familiar to Explorer users.



Metadata (new Grid data type)

- Extends existing data formats to include critical descriptive data
- Enables proper interpretation of data throughout the processing stream
- Includes:
 - labels for data fields, dimensions, and units
 - geographic area and map projection data
 - missing data values
 - identification tags for geometries
- Modules implemented:
 - readhdf2g, contour2g, isosurface2g, orthoslice2g, wireframeg, HorizGridSlice, VertGridSlice, SliceTime, GemDraw Map, GemGetGrid, Gem Topo, GemGridFunction, GridAxisVertical, GridMapping, GridTitle, PrintGridMetadata.

Data Acquisition

HDF file reader

- improved UI
- multiple lattice outputs
- full N-d support for input/output
- automatic backward compatibility with older HDF libraries
- support all 3 Explorer coordinate types
- provide detailed information for file contents (e.g. calculate min,max)
- relevant module : ReadHDF2

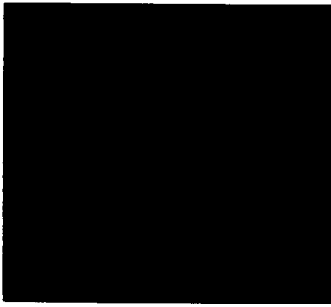
HDF image/palette reader

- relevant modules: Read_HDF_Palette, Read_HDF_Image

DTM connection

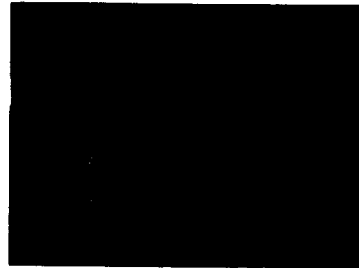
- get data from processes on remote machines
- relevant modules: Read3DDTM

Visual Idioms



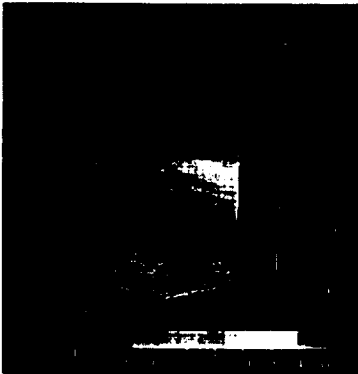
Contour

- capability to set contour interval in addition to # of contour levels
- controls adjusted to input data range
- show zero contour & dim negatives options
- contour data range can be set fixed or data dependent
- relevant modules: [Contour2](#) , [Contour2g](#)



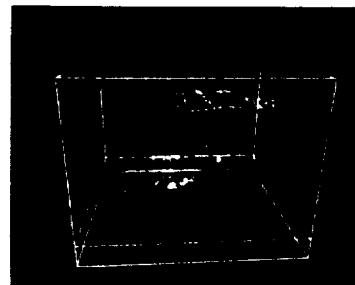
Vector

- different vector styles
- different data to color mapping
- different vector scaling
- controls adjusted to input data range
- relevant modules: [VectorDisp](#)



Annotation

- annotated vertical axis with tick-marks and labels
- geometric representation of a color map
- 2D or 3D titles which incorporate metadata
- relevant modules: [GridAxisVertical](#), [ColorKey](#), [GridTitle](#)



Particle advection

- generates points in a user specified subvolume of input data space
- advects all the points in space for the duration of display interval
- displays time dependent/independent particle trajectories
- relevant modules: [GenPoints](#), [ParticleAdvect](#), [ParticleTrace](#)

Laser Disk Recording



- provides an interface to the Sony laser disk recorder
- record user and title information on the disk
- relevant module : Laser_Disc_Recorder



Prototypes

• Frame flipbook

- an image flipbook for storage and playback of images
- a geometry flipbook for storage and playback of sequences of geometric scenes while retaining 3D interactive control of viewing parameters

• Simplified user interface to Explorer for basic visualization tasks

- provides simple to use quick access to common visualization methods (contours, isosurfaces, etc.)
- generates Explorer maps by clicking on a table of data variables vs visual idioms

• DIALOGUE

- is an object oriented message passing system for peer style communications
- uses DTM as the transport mechanism
- supports synchronous and asynchronous communication
- is used to move data between running models and

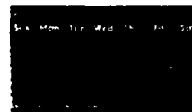


Explorer Alpha and Beta Testing

- Numerous bug reports to Explorer development team



- Contributed to each Explorer release: 1.0, 2.0a, 2.0b1, 2.0b2 and 2.0
- Contributed to the concept of looping and animation in Explorer
- Contributed to the development of the scripting language
- Pushed Explorer to its limits
 - _ Large data sets
 - _ Virtual Reality



Deliverables and Time Table

September 1993:

- beta release of NCSA Pathfinder Explorer modules to NCSA anonymous server
 - SGI binaries, standard module help files
- release of GEMVIS to Unidata
 - module source, documentation, sample data

October 1993:

- beta release of Inventor animation tool to NCSA anonymous server
 - SGI binaries, help file
- release of GEMVIS to COSMIC
 - module source, documentation, sample data

Spring 1994:

- **Workshop on PATHFINDER**
 - held at NCSA
 - other sites connected by video conferencing
-

Summer 1994:

- **final release of NCSA Pathfinder Explorer modules to NCSA anonymous ftp server**
 - updated SGI binaries, source code, user documentation, sample datasets, sample Explorer maps
 - Inventor animation tool source , user's guide, sample data files

Getting Started with NCSA Mosaic

Marc Andreessen
 Software Development Group
 National Center for Supercomputing Applications
 605 E. Springfield, Champaign IL 61820
 marca@ncsa.uiuc.edu

May 8, 1993

1 Introduction

NCSA MOSAIC is a distributed hypermedia system designed for information discovery and retrieval over the global Internet. MOSAIC provides a unified interface to the various protocols, data formats, and information archives used on the Internet and enables powerful new methods for discovering, using, and sharing information.¹

This document is a brief overview of the MOSAIC environment. For more information on any subject covered in this document, please feel free to send an electronic mail message to:

mosaic@ncsa.uiuc.edu

Alternately, feel free to contact the author directly.

2 The Pieces

NCSA MOSAIC uses a client/server model for information distribution - a *server* sits on a machine at an Internet site fulfilling queries sent by MOSAIC *clients*, which may be located anywhere on the Internet.

Units of information sent from servers to clients are simply termed *documents*. Documents may contain plain text, formatted text, inlined graphics, sound, and other multimedia data, and *hyperlinks* to other documents that may be in turn located anywhere on the Internet.

The NCSA MOSAIC for the X Window System client looks like Figure 1: each underlined word or phrase is a *hyperlink*; clicking on it with a mouse button causes the client to connect to the appropriate server anywhere on

¹MOSAIC is based on the World Wide Web technology from CERN in Switzerland and uses the World Wide Web common client library for much of its low-level communications layer.

the Internet and to retrieve and display the referenced document.

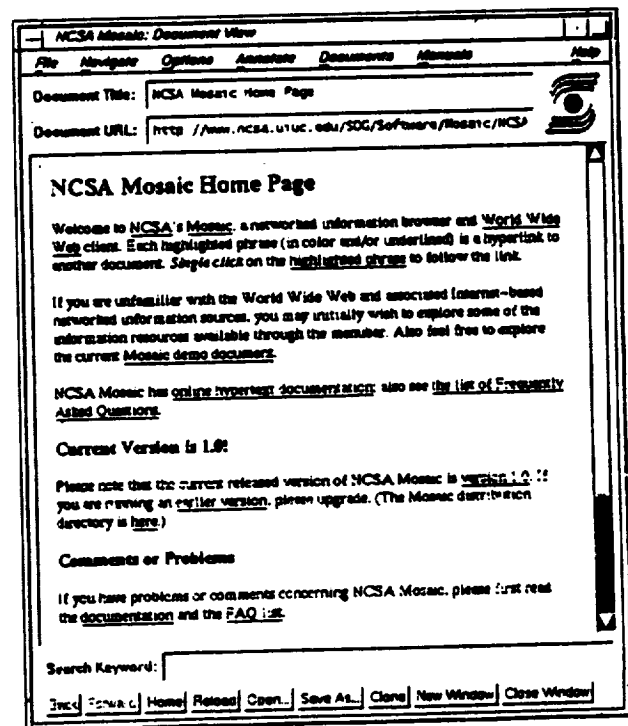


Figure 1: NCSA MOSAIC for X — Home Page

When the MOSAIC environment is used to serve and access a document containing formatted text and inlined graphics on the Internet, the result may look something like Figure 2.

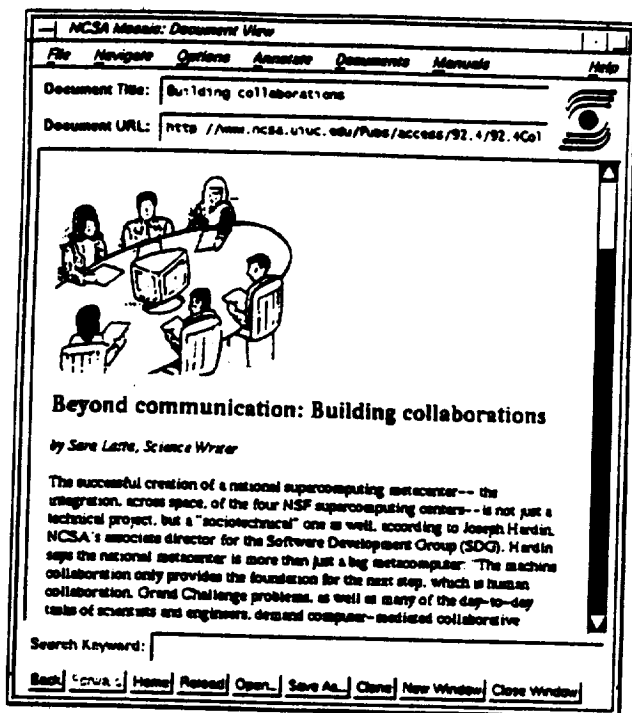


Figure 2: NCSA Mosaic for X — Inlined Image

3 Setting Up The Client

Currently, NCSA has developed and released version 1.0 of the X Window System Mosaic client, which can be used on almost any modern Unix-based graphics workstation (e.g., Sun Sparc, IBM RS/6000, DEC 5000 or Alpha, Silicon Graphics IRIS). The Macintosh client is under development and slated for 1.0 release in summer '93, and work on the Microsoft Windows client will begin in early summer '93. *The remainder of this document currently assumes use of the X client.*

The NCSA Mosaic anonymous FTP site is:

```
ftp.ncsa.uiuc.edu
```

Binary executables of the X Window System client for several common Unix platforms are in this directory:

```
/Mosaic/xmosaic-binaries
```

Source code for the X Window System client is in this directory:

```
/Mosaic/xmosaic-source
```

The NCSA Mosaic Technical Summary, which details Mosaic's current capabilities and our plans for further development in 1993, is in this directory:

```
/Mosaic/mosaic-papers
```

3.1 Basic Client Setup

Pulling down, installing, and running a Mosaic binary is straightforward: there are no configuration files or environment variables that need to be set for the program to work.

Note: If you are running on a Sun workstation under OpenWindows, you may experience a large number of spurious warnings when you first execute the program; these are not serious and will not impair the proper functioning of the program, and the online Frequently Asked Questions list explains how to make them go away.

Note: When Mosaic is executed, the first thing it does is pull the Mosaic "home page" (startup document) over the Internet from an NCSA server. If you get an error message instead of this home page, your connection to the Internet is less than complete. Please either run Mosaic on a system fully connected to the Internet or contact the author for more information.

3.2 Complete Client Setup

To allow interaction with a wide variety of data formats for images, audio, video, and typeset documents currently popular on the Internet, Mosaic relies on a number of *external viewers*: programs separate from Mosaic that are invoked when necessary to display certain types of data.

Here is a list of recommended external viewers for use with Mosaic:

- *xv*: A popular image display utility for X Window System platforms, *xv* can display images encoded in GIF, JPEG, TIFF, and several other formats. *xv* can be obtained via anonymous FTP to:

```
export.lcs.mit.edu
```

The filename of *xv* on that server is currently:

```
/contrib/xv-3.00.tar.Z
```

- *showaudio*: A shell script capable of transparently playing several formats of audio on several platforms, *showaudio* is available as part of the *metamail* multimedia mail toolkit available at the following FTP site:

```
thumper.bellcore.com
```

The current filename for the *metamail* distribution is:

```
/pub/nsb/mm2.4.tar.Z
```

The *metamail* program itself is used by MOSAIC to process multimedia documents that use the MIME format. MIME is not yet widely used as a document format, but it is becoming increasingly popular and is likely to become a *de facto* Internet multimedia document standard in the near future.

- *mpeg_play*: A display utility for MPEG-format animation sequences and video clips, *mpeg_play* is available from the following FTP site:

`postgres.berkeley.edu`

The current filename is:

`/pub/multimedia/mpeg/mpeg_play-2.0.tar.Z`

External viewers useful for viewing typeset documents in TeX's DVI format and in PostScript are harder to compile and install, but here's what we recommend:

- *ghostview*: A user interface to a PostScript viewer called *ghostscript*, *ghostview* offers the best freely available online PostScript viewing technology currently available. The FTP server is:

`uxc.cso.uiuc.edu`

You will need the following files:

`/gnu/ghostscript-2.5.2.tar.Z`
`/gnu/ghostscript-fonts-2.5.2.tar.Z`
`/gnu/ghostview-1.4.1.tar.Z`

- *xdvi*: An online viewer for documents in TeX's DVI format, *xdvi* can be found at:

`export.lcs.mit.edu`

The filename is:

`/contrib/xdvi.tar.Z`

These are the basic external viewers that MOSAIC will assume are present when it encounters multimedia data of the appropriate types. There are other such viewers; more information is available in the online documentation.

4 Setting Up The Server

You can serve documents to NCSA MOSAIC sessions running anywhere on the Internet from an ordinary anonymous FTP server. However, for performance reasons, we recommend setting up a *HyperText Transfer Protocol* (HTTP) server².

²HTTP is a stateless, lightweight, and extremely fast alternative to FTP for network file transfers in the World Wide Web environment.

NCSA has developed a freely available HTTP server. It is currently in late beta and is available at:

`ftp.ncsa.uiuc.edu`

The directory is:

`/Mosaic/ncsa_httpd`

Full configuration instructions are provided with each of the binary and source distributions available there. Feel free to contact the author if you have any questions or problems.

5 Where To Go From Here

Assuming you have successfully installed the MOSAIC client on your system, you should spend some time to become familiar with the interface it presents to the universe of information already available on the Internet. The MOSAIC "demo document", available via a hyperlink in the MOSAIC home page, is a self-contained overview of MOSAIC's capabilities with hyperlinks to a wide variety of interesting information sources.

5.1 Serving Information

If you wish to serve information to the Internet, you should first learn about *HyperText Markup Language* (HTML), the SGML-based markup language used for formatted hypermedia documents in MOSAIC.

You should also learn about the *Uniform Resource Locator* (URL³) scheme for consistently naming documents accessible on the Internet.

At the very end of the MOSAIC demo document (mentioned above), there are hyperlinks to primers on HTML and URL's. Those primers in turn point to more advanced information sources.

6 Acknowledgements

NCSA MOSAIC for X is being implemented by Eric Bina and Marc Andreessen; NCSA MOSAIC for the Mac is being implemented by Aleks Totic. The project lead for NCSA MOSAIC is Joseph Hardin. We are indebted to Tim Berners-Lee and the World Wide Web project at CERN for their vision, ideas, and common client library code, as well as to the many people on the Internet in two dozen countries who are actively providing feedback on NCSA MOSAIC's capabilities.

³The "U" in URL previously stood for Universal.

HANDLING INTELLECTUAL PROPERTY AT UCAR/NCAR

**Dr. William Buzbee
National Center for Atmospheric Research**

August 4, 1993

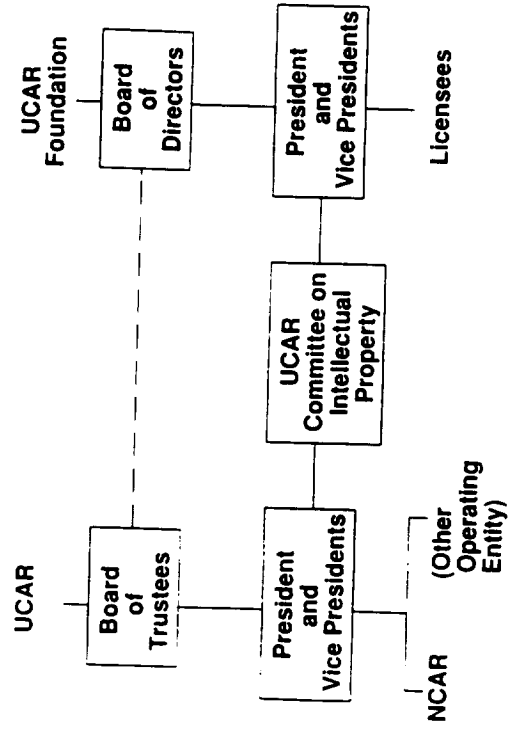
Handling Intellectual Property at UCAR/NCAR

Bill Buzbee

**Presentation to the Applied Information Systems Research Program (AISRP) Workshop III
August 4, 1993**

Enabling Legislation

- President Reagan's Executive Order of April 10, 1987**
- Public Law 98-960, 1984**
- Bayh-Dole Patent and Trademark Amendments of 1980**
- Stevenson-Wydler Act and Federal Technology Transfer Act, 1986**



The Intellectual Property Process for Software

Developer discloses Intellectual Property

UCAR Intellectual Property Committee assesses market potential

If "no market," ownership assigned to developer or public domain

If "market exists"

1. UCAR copyrights/patents
2. NSF obtains nonexclusive and royalty-free license for government purposes
3. Dollar award to employees
4. UCAR Foundation seeks licensee

NCAR Scientific Computing Division

UCAR Licenses

Value Added Remarketer --- negotiable

User License

Includes maintenance

Includes user assistance via e-mail and phone

Installation

Usage

Bug reports

User talk-group (ucarg-talk@ncar.ucar.edu)

Universities

Below cost

Site-wide

Others

Cost recovery based on number of users

NCAR Scientific Computing Division

An Effective Approach to Technology Transfer

Protection

For VAR via license

For U.S. Intellectual Property

NCAR Scientific Computing Division

NCAR Graphics

What's New in 3.2?

by

Bob Lackman

Scientific Visualization Group

SCIENTIFIC VISUALIZATION GROUP, 510
MATERIALS CENTER FOR ADMINISTRATION, B1, MARTIN

Functionality

- A C language binding
- All utility entries
- ANSI GKS C binding
- Raster contouring
- High quality filled fonts
- New Documentation
- man pages
- Fundamentals Document
- Contouring and mapping tutorial

SCIENTIFIC VISUALIZATION GROUP, INC.
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

Functionality (Cont.)

- Utility upgrades
- Contouring
 - Isosurfaces
 - Streamlines
 - Vectors
 - Mapping
 - Text

SCIENTIFIC VISUALIZATION GROUP, INC.
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

Functionality (Cont.)

- New raster utilities
- rascal
 - Concatenate images
 - Format conversions
 - Sun, xwd, NRIF, AVS, HDF, SGI, Abekas
 - Resampling
 - rasls
 - List files with resolution, encoding, and frame count
 - rasview
 - rassplit

SCIENTIFIC VISUALIZATION GROUP, INC.
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

Functionality (Cont.)

- New examples
- ncaregex <example name>
 - Many new examples
 - Use "man ncaregex" to see a list
 - Can generate all examples of a utility
 - ncaregex <example name>
 - New examples using the C bindings
 - Use "man ncaregex" to see a list

SCIENTIFIC VISUALIZATION GROUP, INC.
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

Functionality (Cont.)

- CGM viewing and output
 - Interactive Display Tool (idt)
 - Recursive zooming added
 - Direct animation added
 - CGM Translators (ctrans/ictrans)
 - Options for zooming and panning
 - Direct NCGM to raster conversion
 - Direct NCGM to PostScript conversion
 - HP LaserJet PCL device added
 - Color palette override for NCGM

SCIENTIFIC VISUALIZATION GROUP, ICT
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH



Unsupported Software

- Explorer modules
 - XY plots
 - Contouring
 - Mapping
- High Level Utilities
 - Examples
 - nhlex <example name>
 - man pages
 - Draft User Guide

SCIENTIFIC VISUALIZATION GROUP, ICT
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH



Version 4.0: Late 1994

- High Level Utilities (HLUs)
 - XY plots, contouring, maps, ...
- The NCAR Command Line interpreter (NCL)
 - Data ingest
 - Algebra (limited set)
 - HLU invocation and resource setting
- NCL scripting language with loops and conditionals
- A prototype GUI
- User Guide, Tutorial, and Ref Manual

SCIENTIFIC VISUALIZATION GROUP, ICT
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH



McIDAS-eXplorer: A VERSION OF McIDAS-X FOR PLANETARY APPLICATIONS

Dr. Sanjay Limaye
University of Wisconsin

August 4, 1993

McIDAS-eXplorer

A VERSION OF McIDAS-X FOR PLANETARY APPLICATIONS
APPLIED INFORMATION SYSTEMS RESEARCH PROGRAM / NASA

S.S. LIMAYE
AND
R. J. KRAUSS, D. SANTEK, P. FRY AND E. WRIGHT, L. SHOMOVSKY AND R. S. SAUNDERS (PL)

OBJECTIVE:

- ADAPT THE McIDAS ENVIRONMENT FOR ANALYSIS OF PLANETARY DATA ACQUIRED BY NASA'S MISSIONS TO SOLAR SYSTEM OBJECTS FOR USE ON COMMONLY USED SMALL COMPUTERS/WORKSTATIONS.
- USE A UNIFIED APPROACH TO ANALYSIS OF DATA ON DIFFERENT SOLAR SYSTEM OBJECTS.
- TOOLS FOR NAVIGATION, CALIBRATION AND VISUALIZATION AND ANIMATION OF SOLAR SYSTEM DATA
- USEFUL FOR RESEARCH, OPERATIONAL AND EDUCATIONAL USE.
- USER EXTENSIBLE DYNAMICALLY



CHALLENGES:

DEVELOP A SYSTEM WHICH HAS:

- UNIFIED APPROACH TO ANALYSIS OF DATA ON DIFFERENT SOLAR SYSTEM TARGETS
- DATA MANAGEMENT BY THE ANALYSIS SYSTEM
- FLEXIBLE SOFTWARE TO ADAPT TO CHANGING NEEDS
- ADAPTABLE USER INTERFACE
- PROVIDE PROCESSING HISTORY FOR ANALYZED DATA

DISTRIBUTION AND SUPPORT OF McIDAS-EXPLORER

- REQUIRE A LICENSED COPY OF McIDAS-X
- NO McIDAS USER GROUP SUPPORT (MUG) FOR McIDAS-EXPLORER COMMANDS
- DISTRIBUTION AND SUPPORT MECHANISM TO BE DEVELOPED



IMPLEMENTATION:

HARDWARE

- IBM RS-6000/320H RUNNING AIX
- SILICON GRAPHICS WORKSTATIONS WITH 24-BIT SUPPORT
- SUN WORKSTATIONS
- HP WORKSTATIONS
- INTEL 386/486/PENTIUM PLATFORMS RUNNING OS/2 AND UNIXWARE

SOFTWARE

- DEVELOP APPLICATIONS WITHIN THE CORE McIDAS ENVIRONMENT SPECIFIC TO THE PLANETARY DATA
- EXPAND/ENHANCE CORE McIDAS AS DEEMED NECESSARY OR DESIRED
 - SUPPORT FOR 24-BIT COLOR
 - EXPANDED DATA DIRECTORY STRUCTURE
 - TOP LEVEL IMPLEMENTATION OF USER INTERFACE FOR NOVICE USERS
 - KEYWORD VALIDITY CHECKING, DEFAULT VALUE COMMUNICATION TO THE USER
- ADAPT 'CORE' APPLICATIONS FOR SOLAR SYSTEM TARGETS
- USE NAIF/SPICE SOFTWARE LIBRARY FOR NAVIGATION AND EPHEMERIS CALCULATIONS AS MUCH AS POSSIBLE
- PRODUCE CODE THAT IS EASY TO READ (BY HUMANS), UNDERSTAND AND MAINTAIN!



FULL-LOW SOFTWARE STANDARD ESTABLISHED BY JPL/NAIF

- ELIMINATE COMMON BLOCKS
- FORTRAN OR C
- EXPLICIT DOCUMENTATION OF INPUT/OUTPUT AND SUPPLEMENTARY NOTES OF KEY ASSUMPTIONS, PROCEDURES ETC.
- USE VERTICAL ALIGNMENT AND SPACES TO MAKE CODE MORE EASY TO READ BY HUMANS AND EASY TO MAINTAIN AND MIGRATE TO DIFFERENT PLATFORMS.
- MODULAR DEVELOPMENT OF APPLICATIONS SOFTWARE
- A MOTIVATED USER SHOULD BE ABLE TO DEVELOP APPLICATIONS CUSTOMIZED TO SPECIAL NEEDS



4

DOCUMENTATION

- IMPROVED ON-LINE HELP
- CONTEXT SENSITIVE HELP
- IMPROVED PRE-SCREENING OF COMMAND PARAMETERS FOR VALIDITY
- IMPROVED ERROR REPORTING AND CORRECTIVE ACTION MESSAGES
- USER GUIDE FOR THE NOVICE USER WHICH WILL INCLUDE A CLEAR DESCRIPTION OF PROCEDURES USED AND LIMITATIONS OF THE COMMANDS AND EXAMPLES
- SUMMARY OF MISSION SPECIFIC DATA AND COMMANDS
- DOCUMENTED SUBROUTINE LIBRARY
- PROCESSING MODULE SKELETONS FOR CUSTOMIZATION BY USERS



5

PROVIDE DISPLAY AND ANALYSIS SUPPORT FOR:

- PLANETARY DATA SYSTEM CD-ROM VOLUMES FOR SOLAR SYSTEM DATA
- EARTH BASED AND HUBBLE SPACE TELESCOPE IMAGES OF SOLAR SYSTEM TARGETS
- NO REAL-TIME INGESTS AT PRESENT BUT COULD SUPPORT ACTIVE MISSIONS (E.G. GALILEO AND MARS OBSERVER) OFF-LINE
 - MULTISPECTRAL IMAGES, ATMOSPHERIC VERTICAL PROFILES, SURFACE DATA, MODEL OUTPUT
- IMAGES
 - VOYAGER 1 AND 2 IMAGES OF THE GIANT PLANETS AND SATELLITES
 - MAGELLAN SYNTHETIC APERTURE RADAR REFLECTIVITY IMAGES OF VENUS
- SURFACE
 - MAGELLAN RADAR ALTIMETER DATA
 - VIKING 1 AND VIKING 2 ORBITER IMAGES OF MARS
 - MARS DIGITAL IMAGE MODEL (MDIM'S)
 - PIONEER VENUS ORBITER CLOUD PHOTOPOLARIMETER IMAGES
 - HUBBLE SPACE TELESCOPE WF-PC IMAGES OF PLANETS
 - GALILEO ORBITER SSI IMAGES OF VENUS, EARTH, GASGRA
 - MARS OBSERVER?
 - EARTH BASED TELESCOPIC IMAGES
 - DATA FROM TERRESTRIAL METEOROLOGICAL SATELLITES



6

SPECTRA

- VOYAGER 1 AND 2 IRIS OBSERVATIONS OF PLANETS AND SATELLITES
- MARINER 9 IRIS OBSERVATIONS OF MARS
- GALILEO NIMS (?)

TEMPERATURE PROFILES

- PIONEER VENUS RADIO OCCULTATIONS
- MARINER 9 MARS RADIO OCCULTATIONS
- VIKING 1 AND 2 MARS RADIO OCCULTATIONS
- MARS OBSERVER?

CAPABLE OF FUTURE SUPPORT FOR DATA FROM NEW MISSIONS

- MARS OBSERVER
- CLEMENTINE
- GALILEO
- CASSINI
- EOS



7

**EXPERIMENTER'S LABORATORY FOR
VISUALIZED INTERACTIVE SCIENCE (ELVIS)**

**Elaine R. Hansen
Colorado Space Grant Consortium
University of Colorado**

August 4, 1993



Elaine R. Hansen

Colorado Space Grant Consortium
Laboratory for Atmospheric and Space Physics
National Center for Atmospheric Research

Goddard Space Flight Center
University of California at Santa Barbara

— PRESENTATION OUTLINE —

- The Players
- Project Overview
- Major Accomplishments
- Plans
- TECHNOLOGY TRANSFER



Elaine R. Hansen

Colorado Space Grant Consortium
Laboratory for Atmospheric and Space Physics
National Center for Atmospheric Research

Goddard Space Flight Center
University of California at Santa Barbara

**EXPERIMENTER'S LABORATORY
FOR
VISUALIZED INTERACTIVE SCIENCE
(ELVIS)**

**Applied Information System Research Workshop
August 4, 1993
Boulder, Colorado**



Elaine R. Hansen

Colorado Space Grant Consortium
Laboratory for Atmospheric and Space Physics
National Center for Atmospheric Research

Goddard Space Flight Center
University of California at Santa Barbara

Experimenter's Laboratory for Visualized Interactive Science

— THE PLAYERS —

• A Group Effort

- Colorado Space Grant Consortium:
 - Elaine Hansen
 - Allison Kipple
 - Mike Portert
- Laboratory for Atmospheric and Space Physics:
 - Margi Klomp
 - Eric Hillis
 - Ed Loehr
- National Center for Atmospheric Research:
 - Joseph Klomp
 - Bill Boyd
 - Scott Davis
- Goddard Space Flight Center:
 - Mari Szczer
 - Jeff Star
- Jet Propulsion Laboratory:
 - Thomas Handley

Colorado Space Grant Consortium
Laboratory for Atmospheric and Space Physics
National Center for Atmospheric Research

Goddard Space Flight Center
University of California at Santa Barbara

Elaine R. Hansen

• **With a Large Support Group**

- Steve Lee (LASP / PDS)
- Alex Goetz (CIRES)
- O. R. White (NCAR)
- Gary Heckman (NOAA)
- Greg Scharfer (CIRES / NSIDC)
- Alice Bertini (CASA / ADS)
- Morris Weisman (NCAR)
- Bill Skamrock (NCAR)
- Clayton Lewis (Computer Science)
- Etc.

• **And Many Student Assistants**

- Tony Colaprete (Solar)
- Jennifer Ray (Earth Systems)
- Sr. Design Lab. (Computer Science Department)
- (Ed Loehr, Scott Davis, Ted Crossman, and Bjorn Svensson)

Colorado Space Grant Consortium
 Laboratory for Atmospheric and Space Physics
 National Center for Atmospheric Research



Elaine R. Hansen
 8/1992

— **PROJECT OVERVIEW** —

ELVIS will . . .

- Help scientists of the '90's to interactively visualize the large, complex, and multidimensional data sets that will be produced by our science missions.
- Support scientific research within and across space science disciplines.
- Provide an interactive exploration and analysis tool at an affordable price, on a variety of popular, affordable workstations.
- Display complex structures within 3 dimensional data fields as shaded-surface and volumetric renderings.
- Handle multiple, simultaneous, and diverse data sets (ingest, present, manipulate).
- Capitalize on existing techniques, technologies, tools, and theories.
- Can be easily tailored by the scientists themselves to fit their research problems and display preferences.
- Provide a user interface that is intuitive for and responsive to the needs of the space scientist.

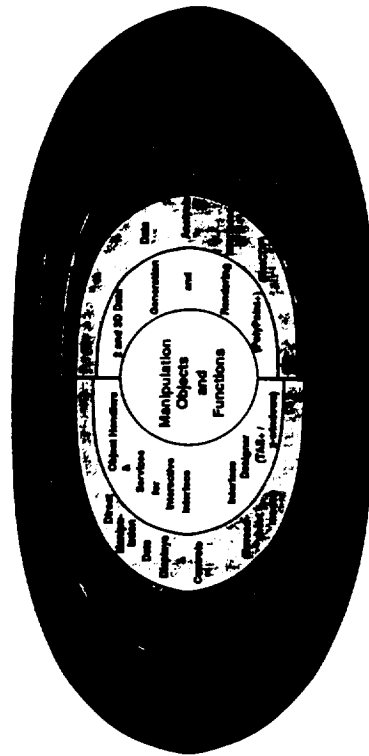
Colorado Space Grant Consortium
 Goddard Space Flight Center
 Laboratory for Atmospheric and Space Physics
 National Center for Atmospheric Research



Elaine R. Hansen
 8/1992

— **ELVIS COMPONENTS** —

ELVIS is the integration of a number of technologies, tools and theories.



Colorado Space Grant Consortium
 Laboratory for Atmospheric and Space Physics
 National Center for Atmospheric Research



Elaine R. Hansen
 8/1992

— **MAJOR ACCOMPLISHMENTS** —

- Applied advanced Human Factors techniques and theories
 - User interviews
 - Cognitive walk-throughs
 - Controlled user tests
 - Prototype evaluations
- PolyPaint upgraded to PolyPaint+
 - Original version restructured to fit into object-oriented context
 - Added new 2D objects
- TAE+ used and augmented
 - Enabled early prototypes of User Interface
 - Color Manager
 - X-Y graphics
- Preparing for the Object-Oriented, direct manipulation user interface
 - Object-oriented design implemented in C++
 - 3D direct manipulation View & Lighting Editor developed
 - Objects implemented to handle data and color
 - Intuitive Color Editor, and Palette developed
- Extensible design to support a scalable set of functions for shaded-surface or volumetric renderings in real or indexed color

Colorado Space Grant Consortium
 Laboratory for Atmospheric and Space Physics
 National Center for Atmospheric Research



Elaine R. Hansen
 8/1992

— MAJOR ACCOMPLISHMENTS —

- **Features Demonstrated Tues-Thurs 3:00-4:00 Include . . .**
 - Interactive User Interface
 - 2D and 3D renderings of some typical space science data sets
 - Direct manipulation 3D View and Lighting Editor
 - Intuitive Color Editor
 - Emphasis on "PolyPaint" rendering package



Blaker R. Hanson
1993

Colorado Space Grant Consortium
Laboratory for Atmospheric and Space Physics
National Center for Atmospheric Research
Guidard Space Flight Center
University of Colorado at Boulder, Boulder

— PLANS —

- **Integration with Data HUB**
- **Integrate the Spreadsheet Engine for:**
 - Building extended interactive visualization capabilities
 - Providing direct manipulation
- **Interface to graphics hardware via Inventor**
- **Integrate visualization of 2D graphics in 3D space**
- **Alpha release in September 1993**
- **User Interviews and User Evaluations with their own data sets**
- **Beta release at end of project**
- **TECHNOLOGY TRANSFER!**



Blaker R. Hanson
1993

Colorado Space Grant Consortium
Laboratory for Atmospheric and Space Physics
National Center for Atmospheric Research
Guidard Space Flight Center
University of Colorado at Boulder, Boulder

— TECHNOLOGY TRANSFER —

- **Technology Transfer has already been occurring in the development phase and will continue through Year 3 of the program.**
- **Technologies transferred in many different ways.**
 1. **Transfer of Advanced Software Tools and Theories**
 - Color Management Enhancements being transferred to GSFC's TAE program and Century Computing for distribution to all TAE+ users.
 - General X-Y Graphing Capability transferred to Century Computing for distribution with TAE+
 - HCI theories transferred to ELVIS project include usability / testability methodology and a direct manipulation, spreadsheet-style user interface concept



Blaker R. Hanson
1993

Colorado Space Grant Consortium
Laboratory for Atmospheric and Space Physics
National Center for Atmospheric Research
Guidard Space Flight Center
University of Colorado at Boulder, Boulder

— TECHNOLOGY TRANSFER — (cont'd.)

2. **Transfer of Intellectual Resources**
 - Students involved as developers, user interviewers, system evaluators, . . .
 - Many additional students exposed to these advanced concepts and capabilities
3. **Technology Leveraging**
 - Taking advantage of Tom Handley's Data HUB capability for data translation into NetCDF
 - In turn, we hope can help shape capabilities of Data HUB
4. **Technology Basis for Future Work**
 - NCAR proposal, with ACTS, for cooperative data analysis
 - AISRP proposals
 - Etc.



Blaker R. Hanson
1993

Colorado Space Grant Consortium
Laboratory for Atmospheric and Space Physics
National Center for Atmospheric Research
Guidard Space Flight Center
University of Colorado at Boulder, Boulder

— **TECHNOLOGY TRANSFER** — (cont'd.)

- 5. Evaluation Testbed**
- Testbed of evolving capabilities
 - For checkout, user demonstration and evaluation
 - Central method of interesting, informing, and training users and of getting ELVIS into the user community
- 6. Spreading the Word**
- Presentation at TAE users meetings
 - Local demonstrations for science meetings, visiting scientists, NASA visitors, etc.
 - A paper (with presentation and demonstration) describing integrated capability
 - Articles on specific technology advances (Item 1)
 - Professional contacts

Colorado Space Grant Consortium
Laboratory for Atmospheric and Space Physics
National Center for Atmospheric Research



Shirley R. Hansen
6/19/93
12

— **TECHNOLOGY TRANSFER** — (cont'd.)

- 7. Technology Insertion**
- Alpha version distributed to 5 selected users in September 1993
 - Beta version distributed, at no cost, at end of project to all interested users
 - Through Software Support Laboratory
 - Will pursue additional funding to maintain, support, and enhance this Beta version to enable us to continue to support the space science user community

Colorado Space Grant Consortium
Laboratory for Atmospheric and Space Physics
National Center for Atmospheric Research



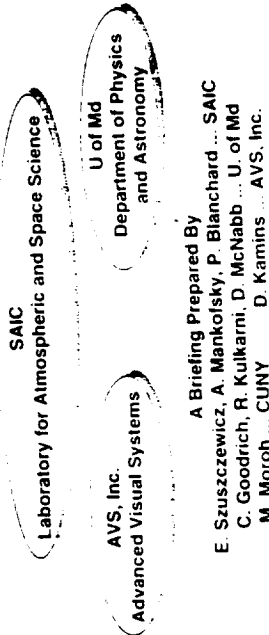
Shirley R. Hansen
6/19/93
13

SAVS: A SPACE AND ATMOSPHERIC VISUALIZATION SCIENCE SYSTEM

Dr. Edward Suszczewicz
 Laboratory for Atmospheric and Space Sciences
 Science Applications International Corporation

August 4, 1993

SAVS A SPACE AND ATMOSPHERIC VISUALIZATION SCIENCE SYSTEM



A Briefing Prepared By
 August 3-5, 1993
 AISRP Workshop, Boulder, Colorado

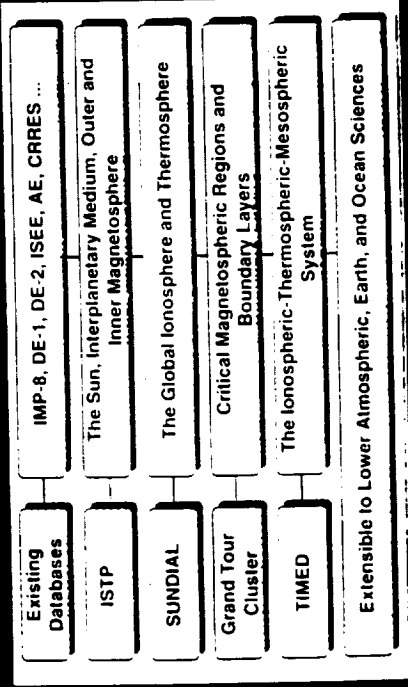
SAVS: DATA ANALYSIS & VISUALIZATION

FOCUS & COMPONENTS

FOCUS
 The Multi-Disciplinary Databases Designed to Understand the Cause-Effect Relationships in the Solar-Terrestrial System and Their Extrapolations to Other Planetary Bodies

MAJOR COMPONENTS
 Widely-Accepted/Commercially-Available Visualization Software (AVS) (Heavily-Leveraged International Users' Module Library)
 Advanced Distributed Database Techniques
 Mathematical, Analytical and Image Processing Tools
 Strongly Developed Sense of Scientific Requirements

A SUBSET OF SPACE SCIENCE APPLICATIONS OF THE SAVS DATA HANDLING AND VISUALIZATION SYSTEM



SAVS

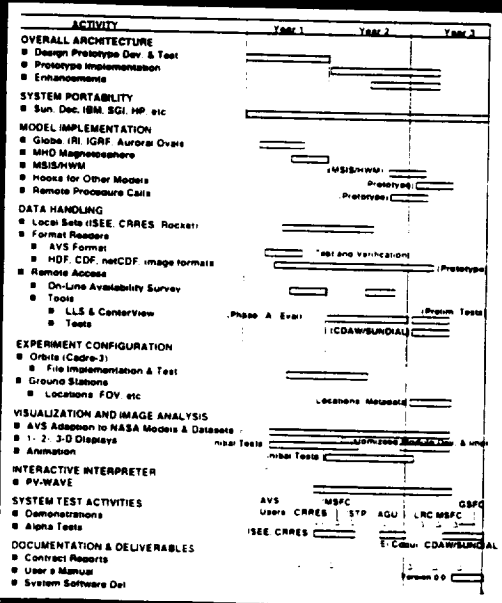
OVERALL OBJECTIVES

- AN INTEGRATED SYSTEM (Ability to Access Data, Models, and Interactively Analyze and Interpret the Findings with Advanced Math, Graphics, and Visualization Tools)
- EASY TO USE AND EXTENSIBLE (Must Service the Impatient Scientist and the Expert Computer Jockey)
- END-TO-END APPROACH (Applicable to Mission Planning, Instrument Design, and Model-Measurement Comparisons)
- PORTABLE AND FINANCIALLY ACCESSIBLE (DEC, HP, IBM, SUN, SGI ... AVS-Only, ~~PV-WAVE~~)

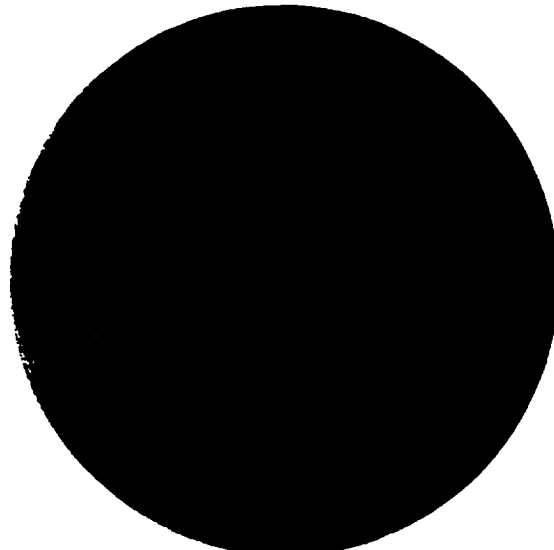
Booklet 4

SAVS

ACCOMPLISHMENTS AND PROJECTED ACTIVITIES



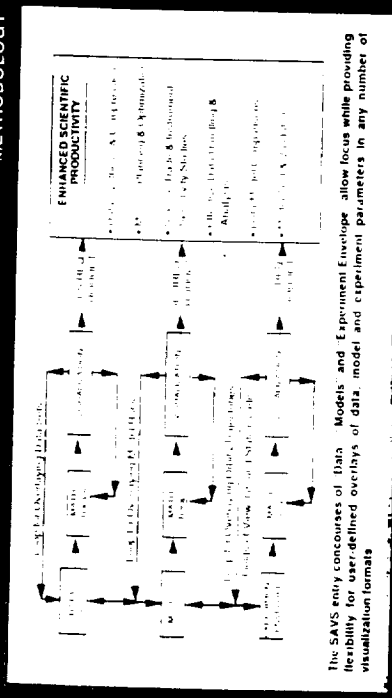
Source: 17



MIMICS SCIENTIFIC METHODOLOGY

INTEGRATED SYSTEM

INTERACTIVE FUNCTIONALITY



Booklet 5

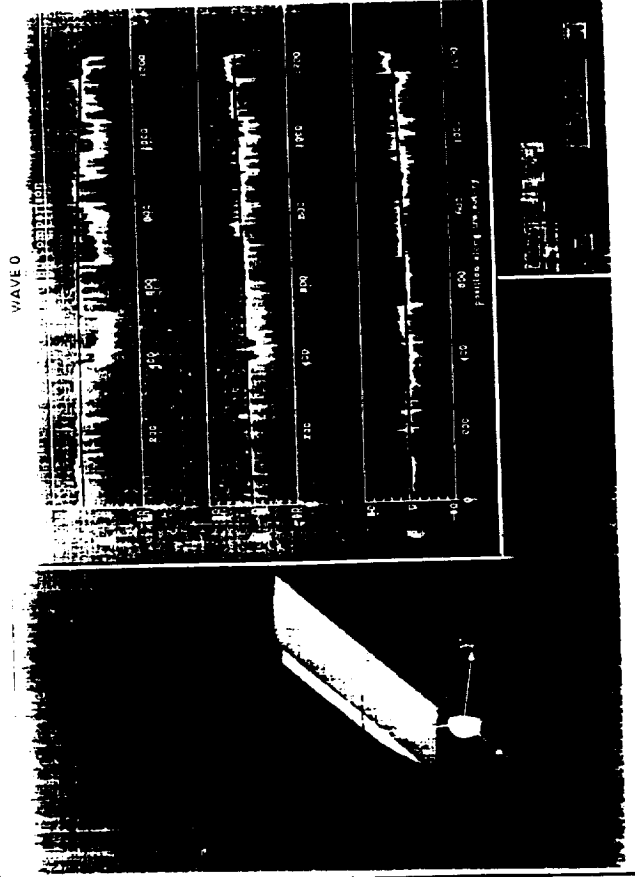
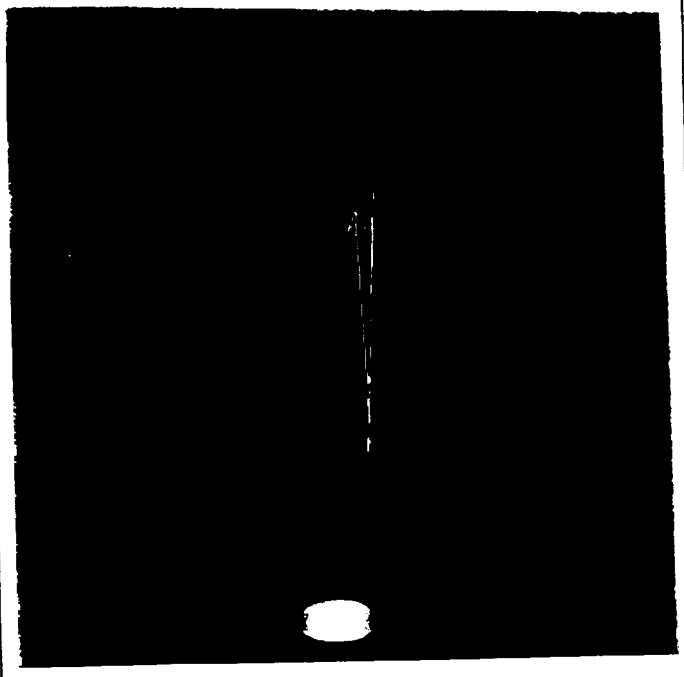
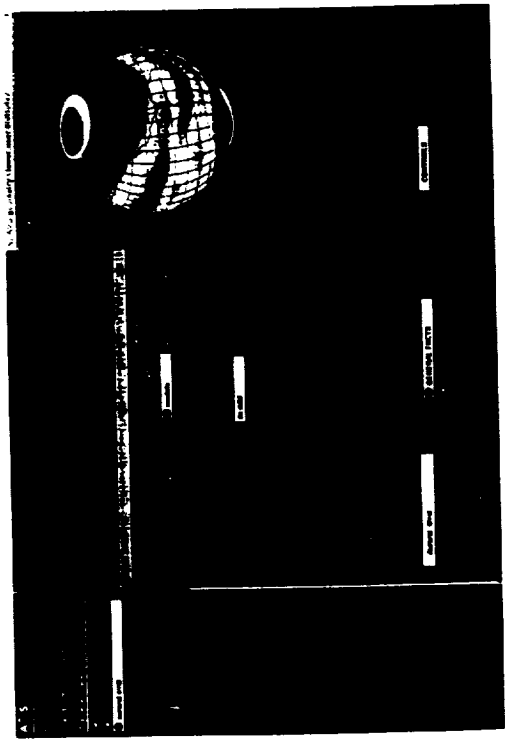
SAVS

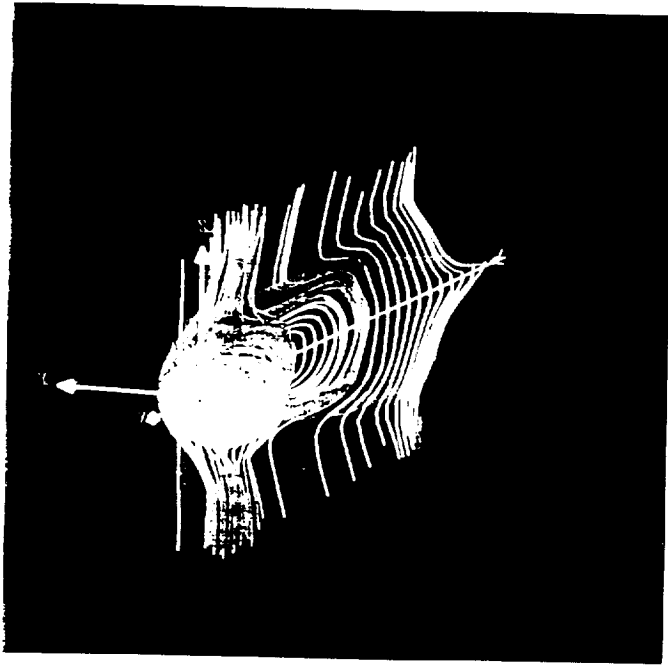
REAL-TIME DEMONSTRATIONS

- New, Easier-To-Use Interactive Front End
- An Application (Visualization is a Process)
 - Orbits + Model + Data
 - Visualization
 - 2-D Stack Plots
 - 3-D Volume Rendering
 - Special Algorithms
 - Coordinate Transformations
 - Field-of-View Projections
- Another Custom Algorithm for NASA Needs
 - 3-D (Data or Model) to 1-D (Orbit Track) Interpolation
- Data Readers and Remote Data Browse and Access
 - HDF and CDF
 - IMP-8, DE-1, DE-2 at NSSDC (Using SAIC CenterView)
- TIMED and ISTEP Mission Planning

Figure 2

3080





SAVS SECOND YEAR ACCOMPLISHMENTS

- Updated and Upgraded Screens, Interfaces and Overall Architecture for Ease, Simplicity, and Extensibility of the SAVS System
- Expanded Utility of General Purpose Model Library (Added MSIS and HWM) and Applied Visualization Techniques to Associated Scalar and Vector Fields
- Developed Customized Modules for Coordinate Transformations and F.O.V Instrument Applications
- Initiated Remote Procedure Call Capabilities
- Initiated Remote Data Access Capability (i.e., Developed Set of Directories, and Staging, Browsing and Extraction Tools)
- Tested Prototype Remote Access and Handling Capabilities on NASA Data (Subsets of IMP: 8, DE-1, and DE-2 Data)
- Developed and Tested SAVS Data Input Modules With Prototype CDF and HDF Readers
- Developed and Tested Generalized Interpolation Module (e.g., Arbitrary 3-D Onto 1-D)
- Developed and Tested SAVS/PV Wave Link to Include Binary Direct Memory Transfer and Bi-Directional Data Transfer

Slide 4



SAVS THIRD YEAR PLANS

- Re-Visit and Freeze Architecture and Interfaces
- Test and Integrate Remote Procedure Call Capabilities
- Extend Development and Test of Prototype Remote Data Access and Handling Capability (CenterView and ILS)
- Extend, Test and Tune CDF, HDF, and NetCDF Readers
- Develop and Test Baseline Image Handling and Remote Sensing Tools
- Establish and Document Recipes for User-Unique Models and Data Formats
- Simulate a Mini CDAA SUNDIAL Campaign
- Fully Integrate and Tune All Customized Modules
- Document and Deliver Software and Write the Users' Manual

Slide 5

LINKWINDS: THE LINKED WINDOWS INTERACTIVE DATA SYSTEM

Dr. Allan Jacobson
Jet Propulsion Laboratory

August 4, 1993

ASJ
7/27/93



The Linked Windows Interactive Data System

Principal Investigator

Allan (Bud) Jacobson - Jet Propulsion Laboratory

Investigator Team

Jet Propulsion Laboratory:
Mark Allen
Ronald Blom
Donald Collins
Lee Elson

San Diego Supercomputer
Center:
Phil Mercurio

Developer Team

Jet Propulsion Laboratory:
Andy Berkin
Martin Orton

San Diego Supercomputer
Center:
Jim McLeod

ASJ
7/27/93



The Linked Windows Interactive Data System

Objectives

1. Develop a software environment to support the rapid prototyping and execution of data analysis/visualization applications.
2. Provide a suite of tools to interactively visualize, explore and analyze large multivariate and multidisciplinary data sets.
3. Develop a user interface which allows maximum data and tools accessibility with a minimum of training.
4. Provide system design and tools to make the environment accessible to application development by users.

ASJ
7/27/93



The Linked Windows Interactive Data System

Key Progress in FY 93

1. LinkWinds version 1.4 released to a significantly expanded number of scientific groups at a variety of institutions.
2. Implemented a realtime data interface to LinkWinds and used it to support the Univ. of Iowa's Plasma Wave Spectrometer aboard the Galileo spacecraft during the Earth 2 encounter in December 1992.
3. New applications added and existing applications expanded (a time-based animator, a 3D scatter plot, overlay capability, macro capability, etc.).
4. Expanded database interface by adding HDF scientific format and GSFC's CDF.
5. Three papers were presented, and published or awaiting publication.

ASJ
7/27/93



LinkWinds



The Linked Windows Interactive Data System

Technology Transfer Approaches

1. Make ingestion of data as easy as possible.
 - a. Provide direct LW interfaces for key datafile formats (HDF, CDF, RAW).
 - b. Arrange to provide DataHub.
2. Provide tools, templates and/or help to enable users to transform their data to acceptable formats.
3. Write special filters to transform unusual data formats in some cases.
4. Provide all new users with an interactive realtime tutorial using MUSE.

ASJ
7/26/93



LinkWinds



The Linked Windows Interactive Data System

Sites Actively Using LinkWinds in Research

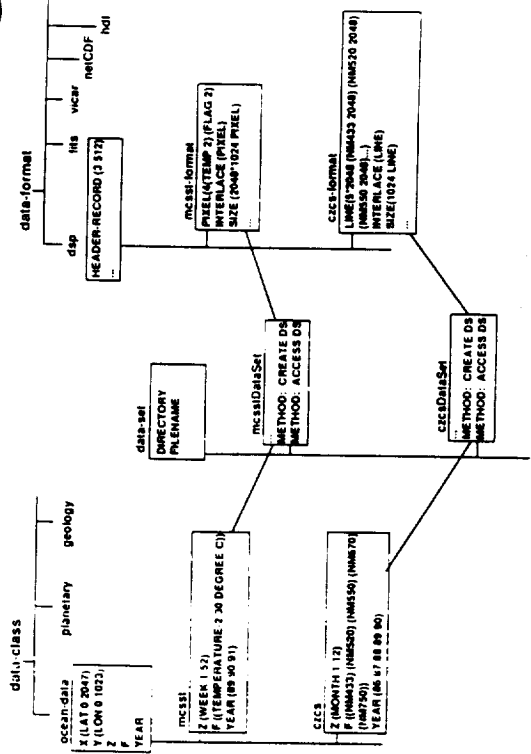
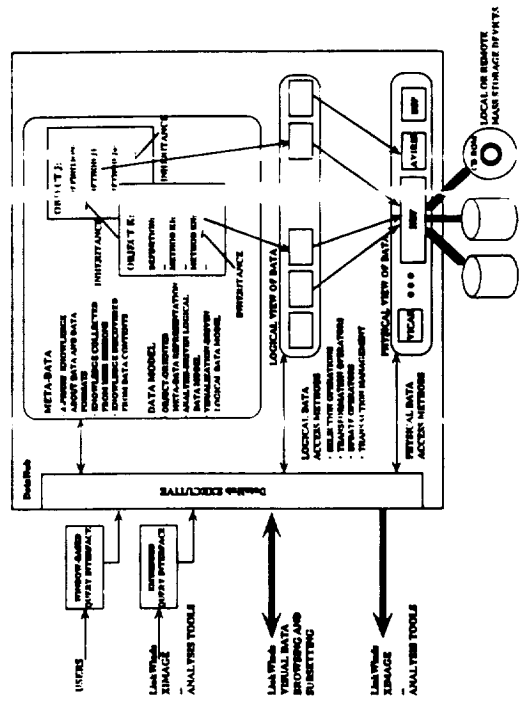
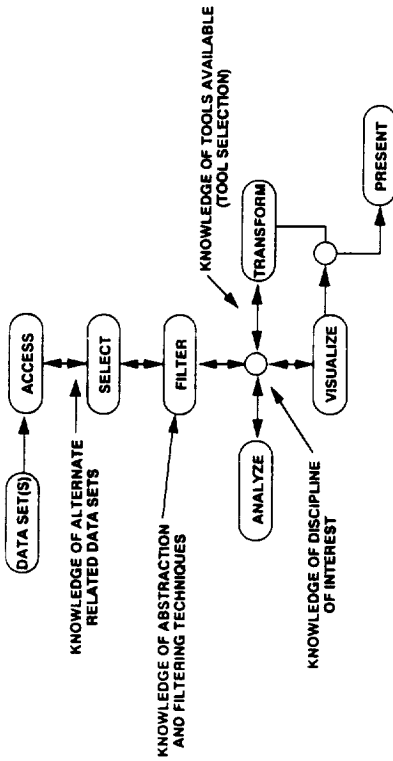
Goddard Space Flight Center
Jet Propulsion Laboratory
Lockheed Palo Alto Research Laboratories
Los Alamos National Laboratory
Louisiana State University
Marshall Space Flight Center
Naval Research Laboratory
Northeastern University
San Diego Supercomputer Center
University of Pittsburgh

ASJ
7/26/93



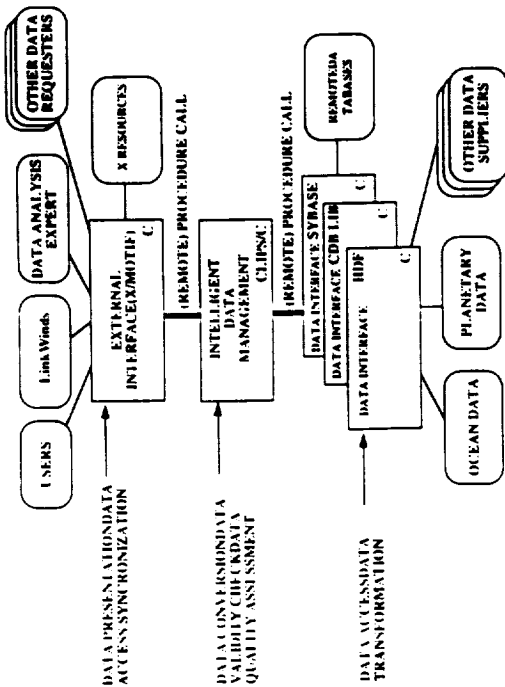
- NEEDS - ADDRESS THE BARRIERS ASSOCIATED WITH DISTRIBUTED, AUTONOMOUS, HETEROGENEOUS SYSTEMS
- DIFFERING ACCESS MECHANISMS
- DATA STRUCTURES
- DATA FORMATS
- DATA SEMANTICS
- INCOMPLETE METADATA

JPL KNOWLEDGE-BASED DATA ANALYSIS AND VISUALIZATION



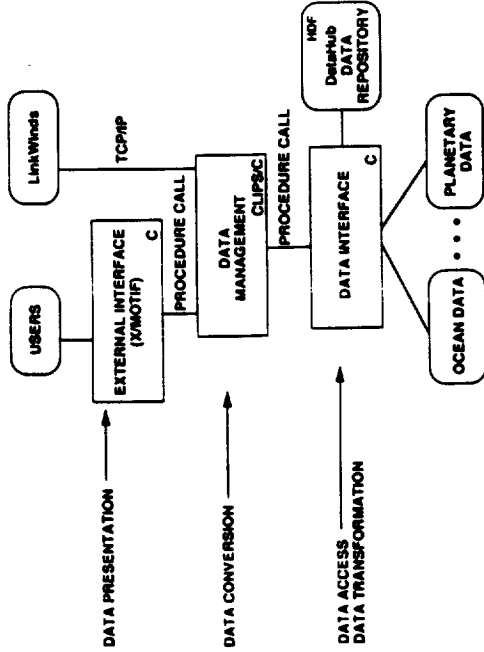


SOFTWARE ARCHITECTURE



JPL

SOFTWARE IMPLEMENTATION



JPL



LESSONS LEARNED

- DEVELOPMENT
 - INITIALLY PLANNED TOOLS WERE: C, FORTRAN, COMMON LISP, PROLOG, SYBASE AND SUNVIEW
 - CURRENTLY: C, FORTRAN, CLIPS, SQL, X WINDOWS, AND MOTIF
 - LITTLE OR NO COST TO USER
 - NOW, MOVING AWAY FROM CLIPS AS OBJECT MANAGER AND THE KERNEL TO C++ WITH SMALL CLIPS RULE-BASED ROUTINES
- EXPERT SYSTEM/MACHINE LEARNING
 - INITIALLY ATTEMPTED ARTIFICIAL NEURAL NETWORK ON CO-INVESTIGATOR'S PROBLEM
 - TOO DIFFICULT TO GENERALIZE THE CAPABILITY TO A GENERIC TOOL
 - NOW, INVESTIGATING MACHINE LEARNING TECHNIQUES FOR FEATURE DETECTION AND LEARNING
- USER INTERFACE
 - INITIALLY, VERY STRUCTURED - LEAD USER STEP-BY-STEP THROUGH PROCESS
 - NOW, "INPUT-PROCESS-OUTPUT" PARADIGM
- SCIENCE PROBLEM
 - INITIALLY, PHYSICAL OCEANOGRAPHY EMPHASIS
 - NOW, PROCESSING AND EXTRACTING INFORMATION FROM MULTI-SPECTRAL DATA

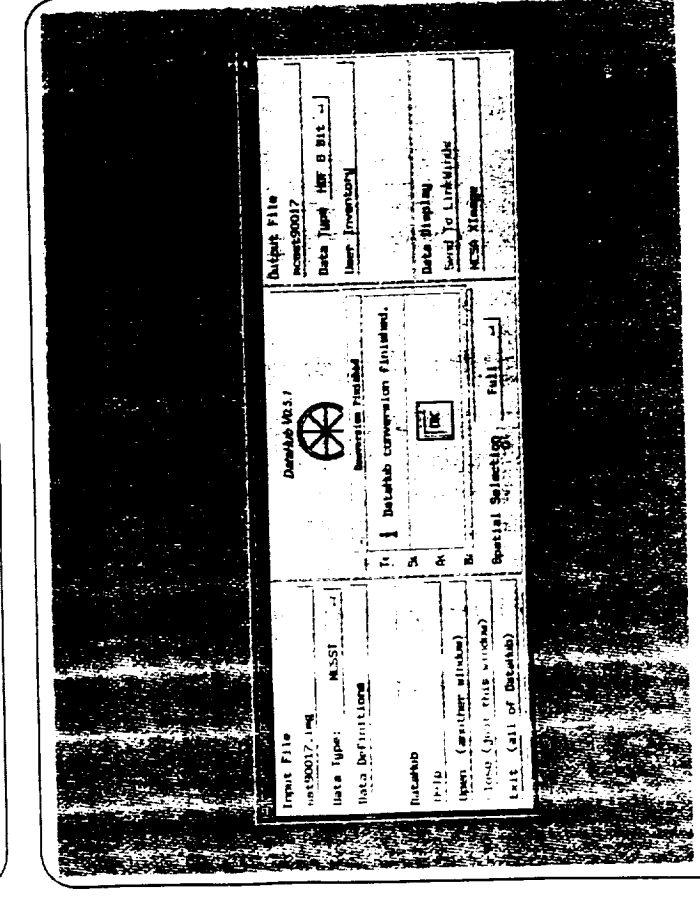
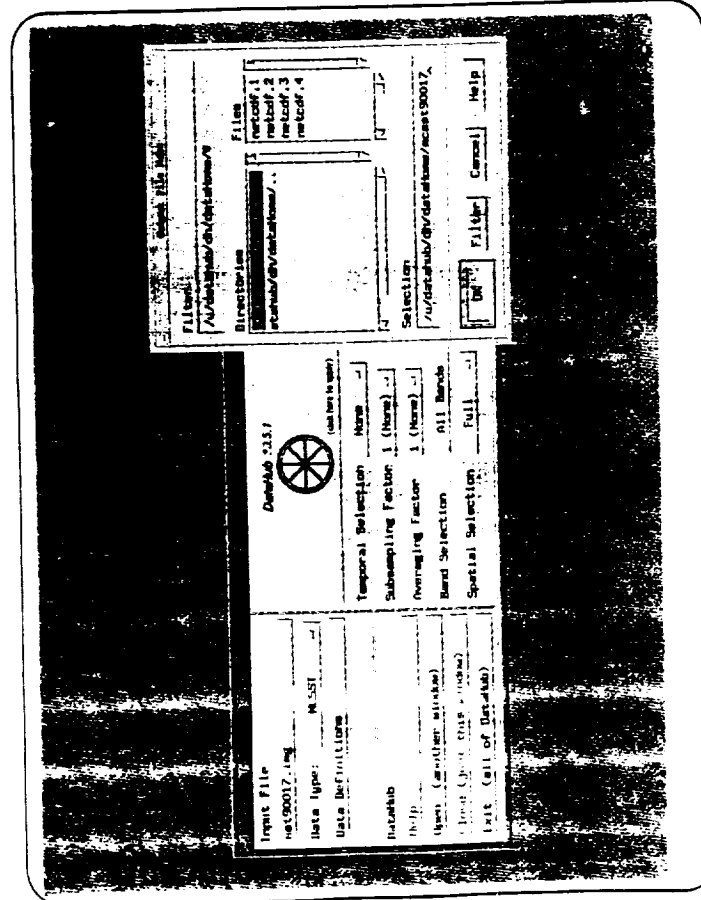
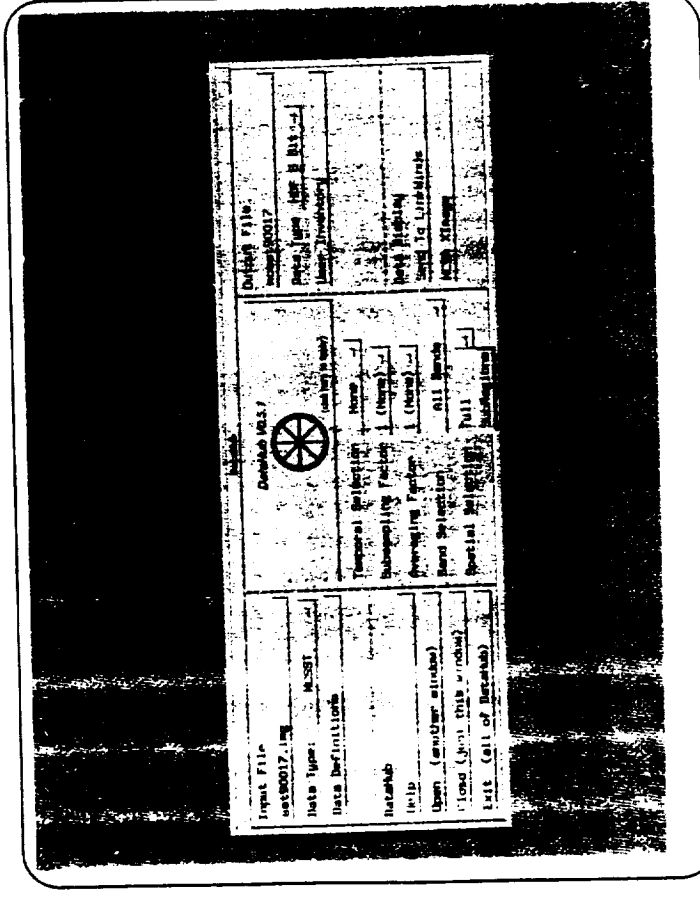
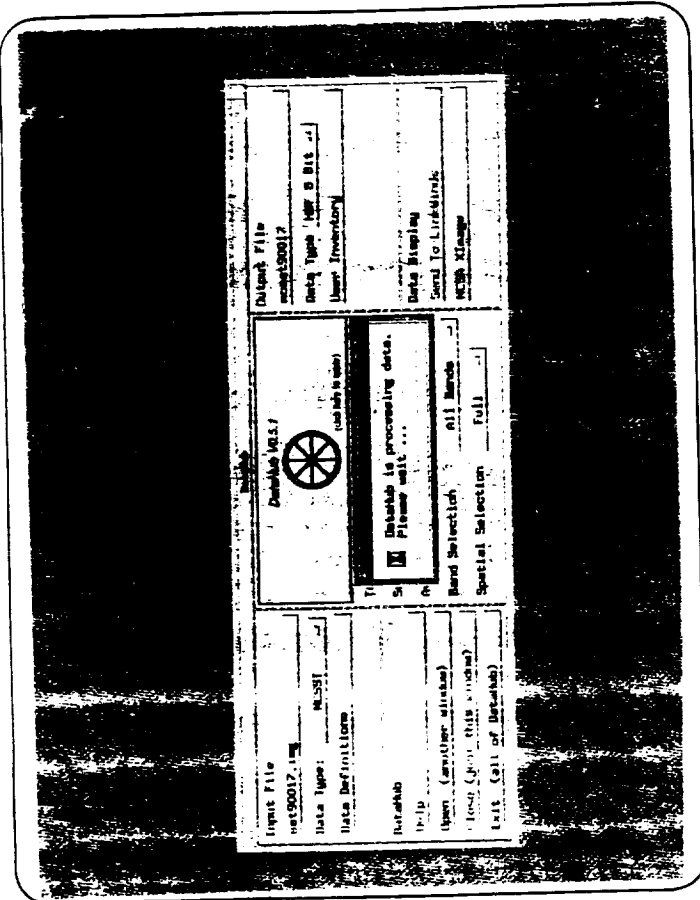
JPL

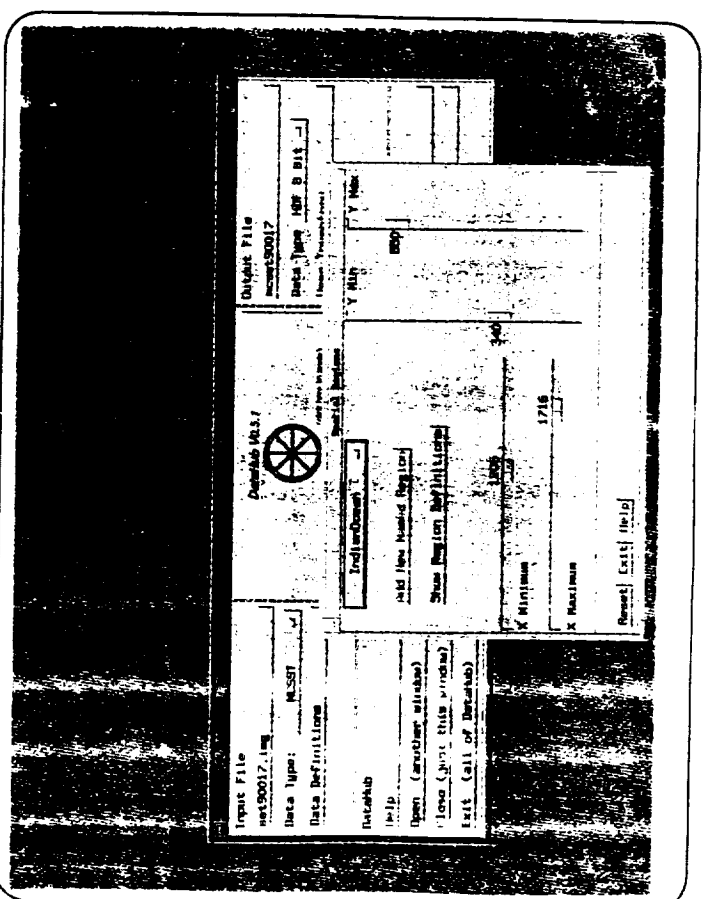
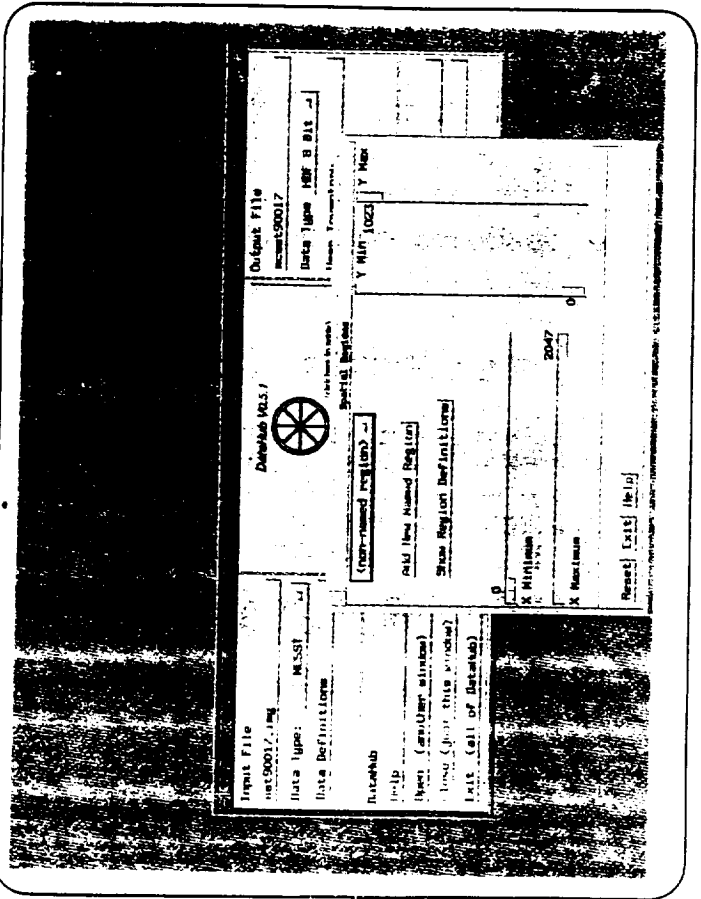
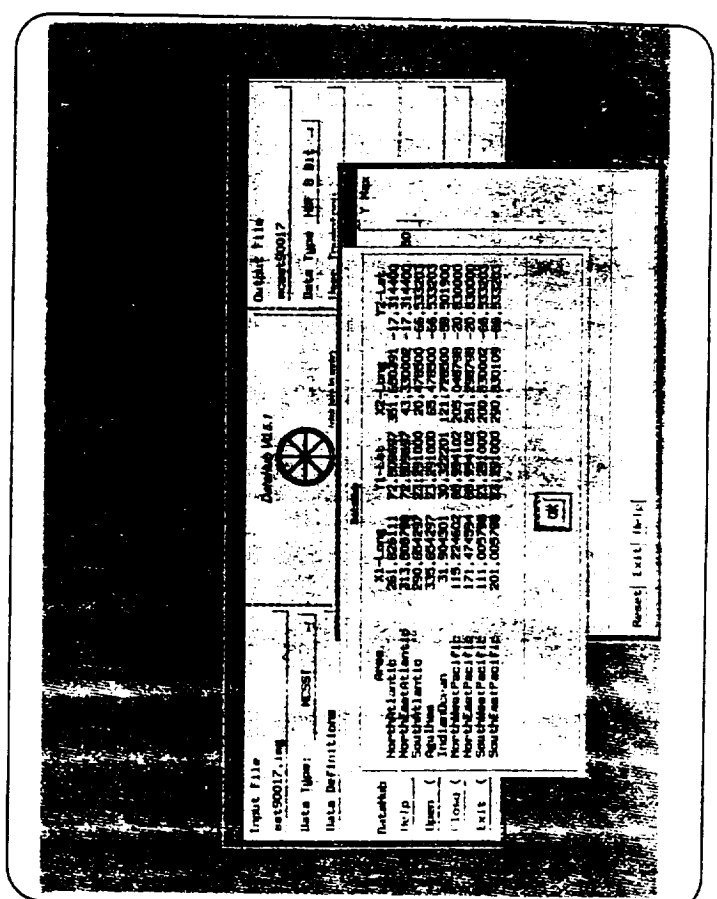
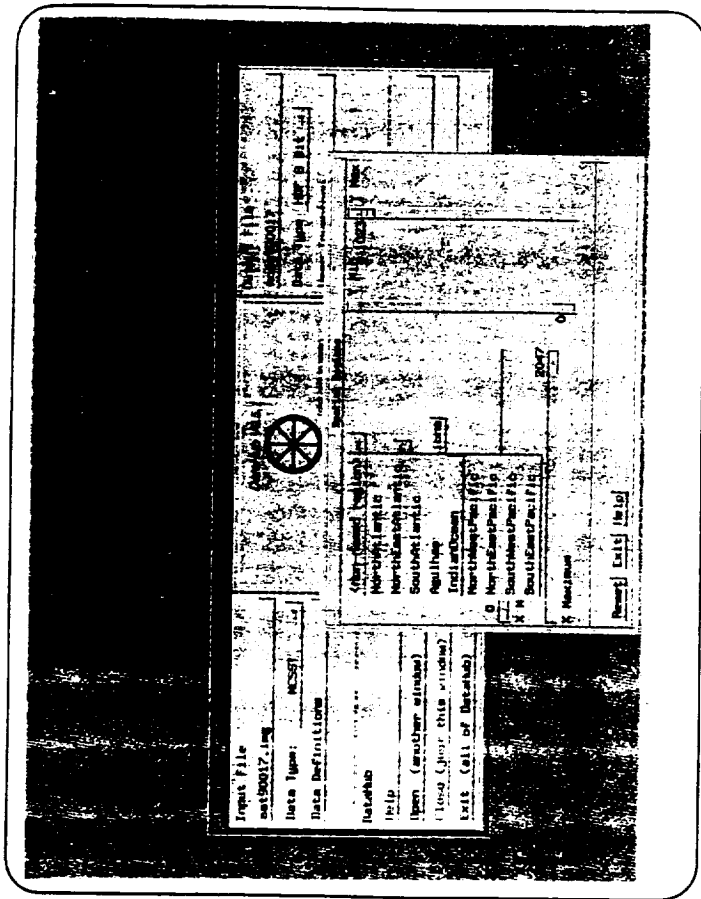


TECHNOLOGY TRANSFER

- USAGL WITH OTHER APPLICATIONS
 - Link Winds
 - DELIVERY WITH Link Winds
 - DATA CONVERSION AND PRE-PROCESSING
 - PRIMARILY TO HIERARCHICAL DATA FORMAT (HDF)
 - POLYPART
 - DATA CONVERSION AND PRE-PROCESSING
 - PRIMARILY TO NETWORK COMMON DATA FORMAT (NETCDF)
 - USAGL FOR SPECIFIC SCIENTIFIC SUPPORT - ROY K. DONIKA (LSU)
 - PRE-PROCESSING, DATA CONVERSION, AND INFORMATION EXTRACTION FROM MULTI-SPECTRAL DATA IN SUPPORT OF TECTONIC HAZARDS RESEARCH

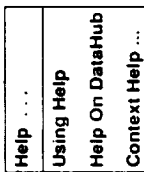
JPL



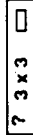




1. USER SELECTS CONTEXT HELP

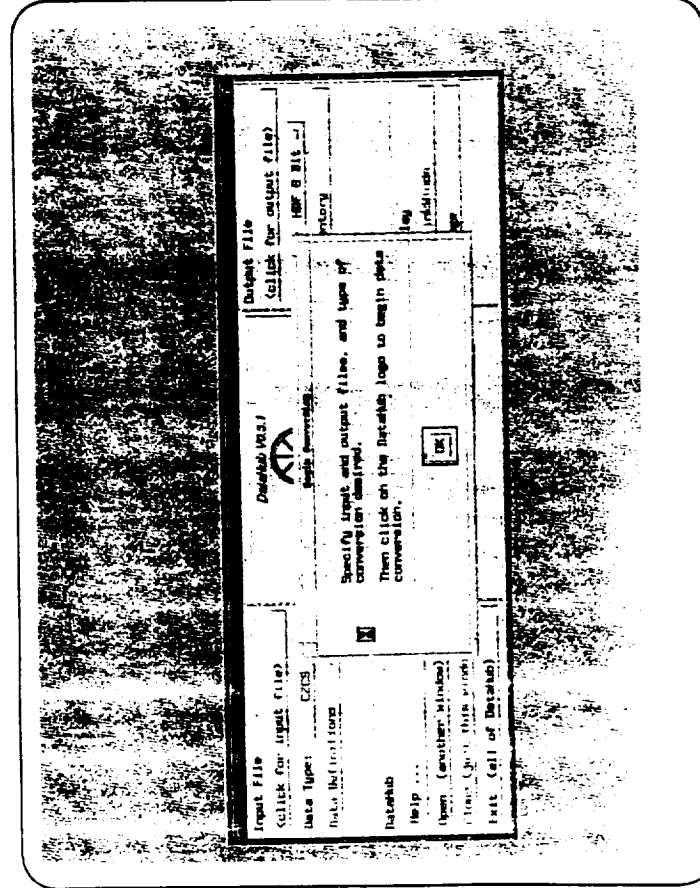
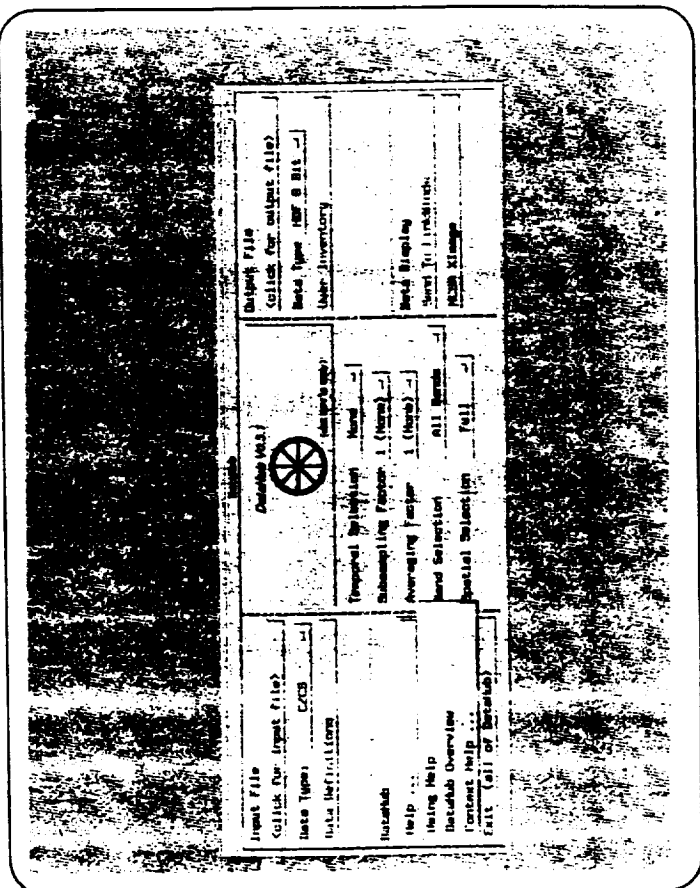
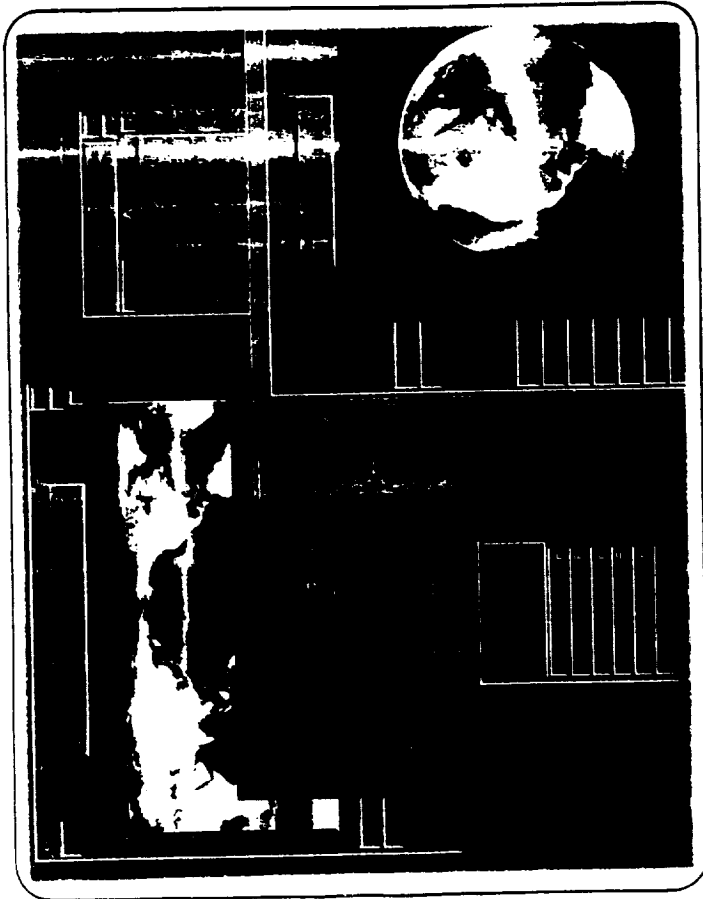
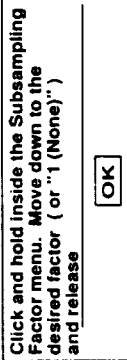


2. DATAHUB CHANGES TO "?" CURSOR



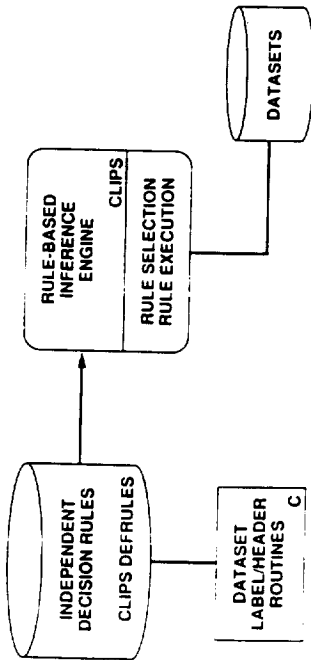
3. USER MOVES CURSOR TO ELEMENT OF INTEREST

4. DATAHUB PRESENTS HELP ONELEMENT



JPL

RULE-BASED CHECKING OF DATASET TYPES



EXAMPLE RULE:

IF THE DATASET HAS A BINARY HEADER,
 AND THE HEADER LOOKS LIKE A HDF DATASET HEADER,
 AND THE HEADER DOES NOT LOOK LIKE A DSP DATASET HEADER,
 AND THE HEADER DOES NOT LOOK LIKE A VOYAGER COMPRESSED
 DATASET HEADER,
 THEN

THE TYPE OF THIS DATASET SHOULD BE HDF.

JPL

FUTURE WORK

ISSUES FOR THE SHORT TERM INCLUDE:

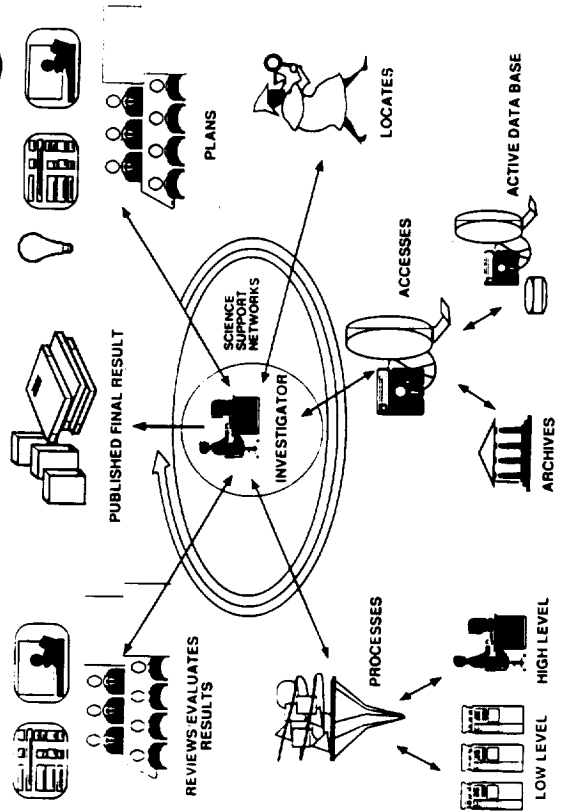
- DataHub-PROVIDED INTERACTION TO MODIFY AND UPDATE RUNTIME RESOURCES,
- A DataHub JOURNAL AND TRANSACTION MANAGER,
- A MORE GENERAL DATA MODEL FOR USER-DEFINED DATA AND CONVERSION,
- AN EXPERT SYSTEM FOR DATASET TYPE CHECKING, AND THE CAPTURING AND UTILIZING OF USER PREFERENCES.

LONG TERM RESET AND IMPLEMENTATION ISSUES INCLUDE:

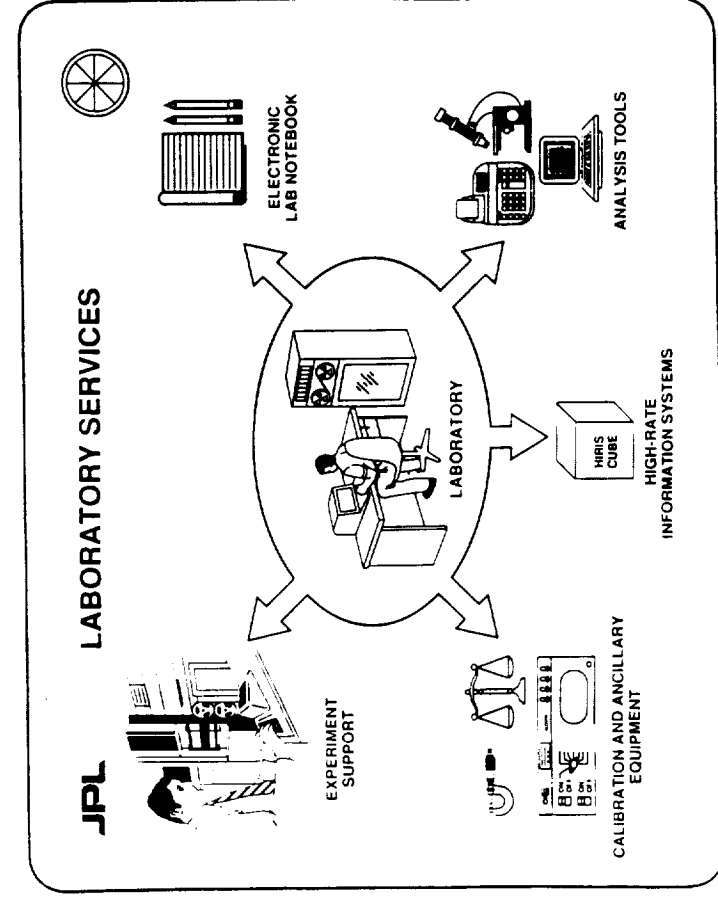
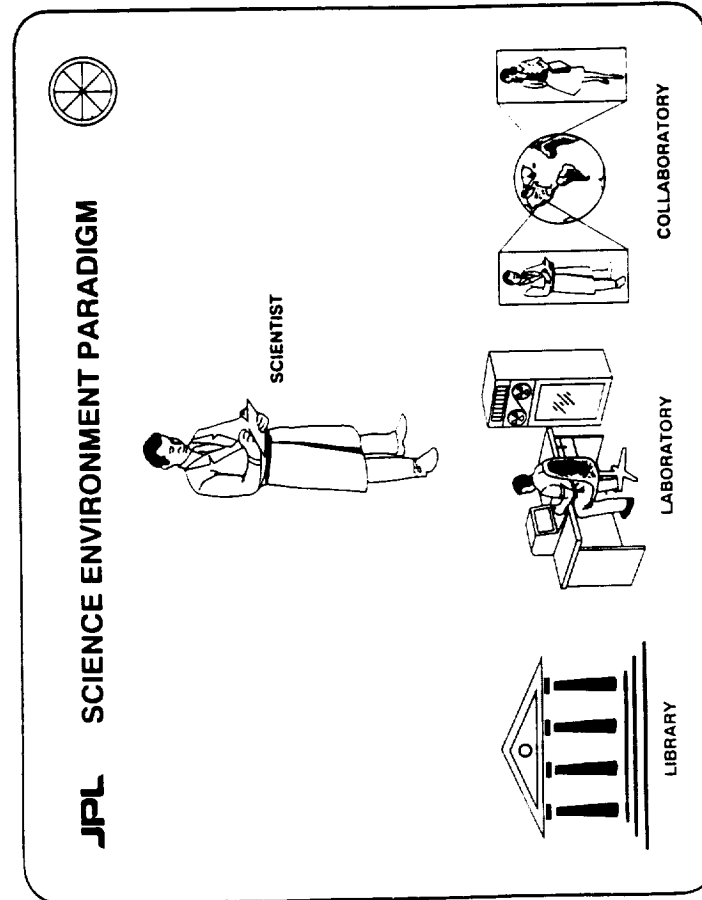
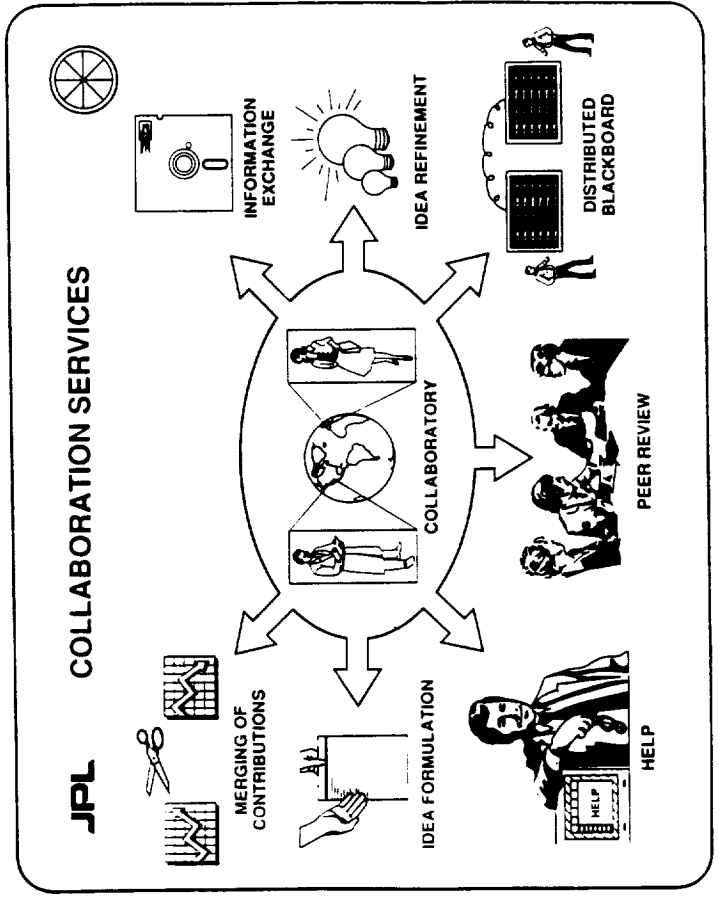
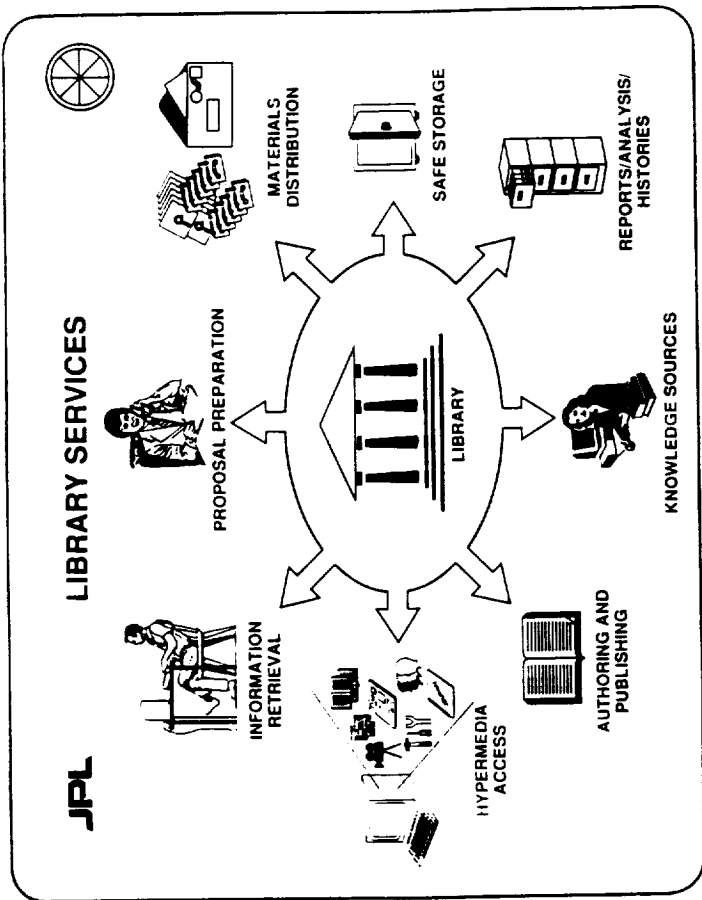
- A DISTRIBUTED DataHub,
- A CONTEXT-BASED AND CONTENT-BASED DATA SEARCH CAPABILITY,
- QUALITY CONTROL AND "CORPORATE" MEMORY OF THE CONVERSION OF INPUT DATA POINTS TO OUTPUT DATA POINTS, AND
- A MORE GENERAL SEARCH PATH AND BROWSING MECHANISM.

JPL

INVESTIGATORS' DISCOVERY CYCLE



BACKUP - VU-GRAPH MATERIAL





JPL SCIENCE ADVISORY PANEL FOR EOS DATA AND INFORMATION

- USER'S VIEW OF THE PERFECT COMPUTING ENVIRONMENT:
"powerful workstations under their control;
painless and affordable access to data from big archives or from other investigators'
workstations;
confidence that the geophysical and biological products they obtain are produced
by knowledgeable people and have established quality and reliability;
communications networks, for easy exchange of information, without the arcane
knowledge currently required to go between systems; and
occasional access to bigger and faster computers."
- TRADITIONAL ROLES OF INVESTIGATORS ARE CHANGING



JPL DATA BASE ISSUES*

DATA BASE MANAGEMENT

SCIENTISTS WANT TO QUERY AND ANALYZE

- RASTER DATA (E.G. SATELLITE IMAGES, DIGITAL ELEVATION GRIDS)
- POLYGONS (E.G. DRAINAGE BASIN BOUNDARIES)
- DIRECTED GRAPHS (E.G. PROFILES OF ATMOSPHERIC TEMPERATURE AND HUMIDITY)
- POINT DATA (E.G. SURFACE METEOROLOGICAL MEASUREMENTS, RIVER DISCHARGE RECORDS)
- TEXT DATA (E.G. ALGORITHM DESCRIPTIONS, PROCESSING HISTORIES)

CHALLENGES

- INDUSTRIAL STRENGTH DBMSS MUST MEET REQUIREMENTS FOR REMOTE SENSING, GEOGRAPHIC INFORMATION SYSTEMS, MORE DATA TYPES, AND OPERATIONS
- CURRENT COMMERCIAL DBMSS ARE NOT GOOD AT MANAGING THESE KINDS OF DATA, WHICH MAY REQUIRE DIFFERENT QUERY INDEXING AND ACCESS METHODS

DATA BASE ISSUES IN EOSDIS, J. DOZIER, 1992 ACM SIG MOD, JUNE 3, 1992



JPL PHYSICAL DATA ORGANIZATION*

ANALYSTS AND MODELERS NEED ACCESS TO

- STORED SATELLITE DATA, DERIVED PRODUCTS, AND MODEL OUTPUT, WHICH ADD UP TO 500 TERABYTES/YEAR
- LARGE, NUMEROUS OBJECTS (E.G. LANDSAT FRAME IS 300 MEGABYTES)
- DATA DISTRIBUTED AMONG 9 ARCHIVE CENTERS

CHALLENGES

- FILE SYSTEMS AND DATABASE MANAGEMENT SYSTEMS MUST UNDERSTAND DISTRIBUTED, TERTIARY MEMORY
- DBMS MUST EFFICIENTLY INDEX AND ACCESS LARGE OBJECTS (INCLUDING CLEVER ABSTRACTS)
- PERFORMANCE MUST BE OPTIMIZED BOTH FOR PRODUCT GENERATION AND USERS' QUERIES

DATA BASE ISSUES IN EOSDIS, J. DOZIER, 1992 ACM SIG MOD, JUNE 3, 1992



JPL REMOTE VISUALIZATION*

RESEARCHERS WANT TO RENDER DATA ON LOCAL WORKSTATIONS, TO

- BROWSE THROUGH MANY IMAGES (COMPRESSED)
- MANAGE USE, AND MANIPULATE LARGE DATA SETS THROUGH DATA BASE MANAGEMENT SYSTEM
- USE "COMPUTATIONAL STEERING" TO GUIDE MODELS

CHALLENGES

- VISUALIZATION SOFTWARE MUST HANDLE LARGE OBJECTS BETTER, THROUGH THE DBMS
- DATA BASE MUST BE QUERIED USING GRAPHS, MAPS, AND IMAGES (AS WELL AS TEXT)
- VISUALIZATION SOFTWARE MUST PROVIDE INTERACTIVE I/O WITH MODELS, SO THAT CHANGES IN MODEL OUTPUT, CAUSED BY INTERACTIVELY CHANGING PARAMETERS, ARE IMMEDIATELY DISPLAYED

DATA BASE ISSUES IN EOSDIS, J. DOZIER, 1992 ACM SIG MOD, JUNE 3, 1992

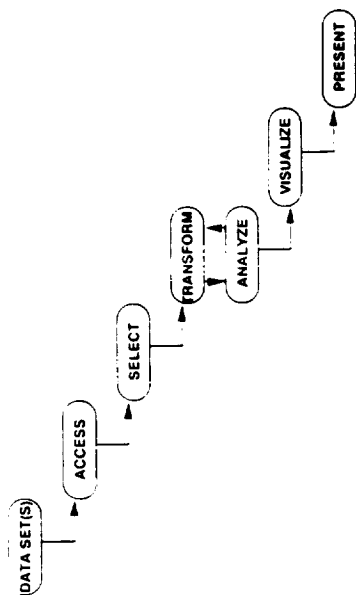


JPL LinkWinds THE LINKED WINDOWS INTERACTIVE DATA SYSTEM

- OF PARTICULAR NOTE IS THE RELATIONSHIP BETWEEN DataHub AND LinkWinds
- LinkWinds – SYSTEM DESCRIPTION
- A VISUAL DATA EXPLORATION/ANALYSIS ENVIRONMENT WITH DATA DISPLAYED IN INTERDEPENDENT WINDOWS. INTERDEPENDENCE IS ESTABLISHED BY "LINKING" VISUALS AND CONTROLS. RESULT IS A GRAPHICAL SPREADSHEET
- A STANDARD GRAPHICAL USER INTERFACE WITH ADDITIONAL DATA-LINKING RULES. RESULTS IN AN INTUITIVE INTERFACE WITH RAPID INTERACTIVITY
- IMPLEMENTED ON AN OBJECT-ORIENTED PROGRAMMING MODEL, WITH "LINKS" ESTABLISHING MESSAGE FLOW PATH. THERE IS AN UNDERLYING COMMAND LANGUAGE (LYNX) BASED UPON SCHEME
- A MULTI-USER SCIENCE ENVIRONMENT (MUSE) REQUIRING A MINIMUM OF NETWORK BANDWIDTH



JPL TRADITIONAL DATA ANALYSIS



JPL MCSST DATA TYPE RESOURCE ENTRY

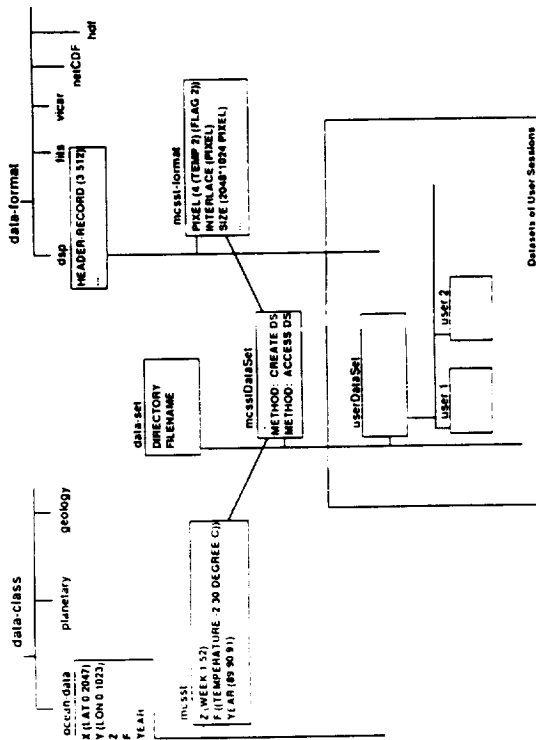
```

MCSST
DSP MCSST
  .resData/mcsst.org
2047
0
1023
0
0
0
2047
1023
0
True
1
1
12
0
  
```

*MCSST typeLabel:
 *MCSST datatype:
 *MCSST directory:
 *MCSST xMax:
 *MCSST xMin:
 *MCSST yMax:
 *MCSST yMin:
 *MCSST x1:
 *MCSST y1:
 *MCSST x2:
 *MCSST y2:
 *MCSST numOIBands:
 *MCSST temporalSelectionSelected:
 *MCSST numOITemporalIntervals:
 *MCSST numOITemporalIntervalsSelected:
 *MCSST numOITYears:
 *MCSST numOITYearsSelected:



JPL USER CREATED DATASETS



JPL REGIONS DEFINITIONS RESOURCES

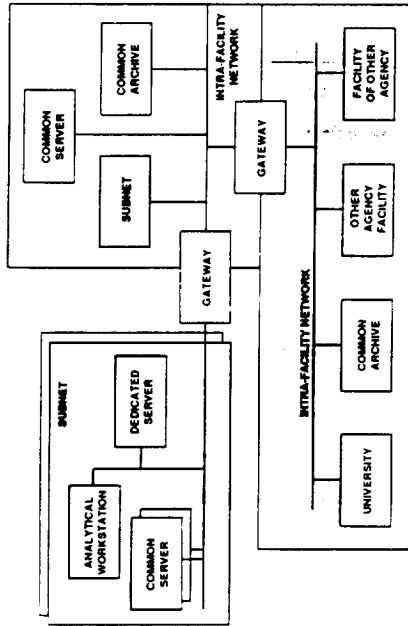
*del.aullRegions:

Region	Lat Begin	Lat End	Long Begin	Long End
*NorthAtlantic:	72.5097	-17.3144	261.8261	351.6504
*NorthEastAtlantic:	72.5097	-17.3144	313.5058	43.3300
*SouthAtlantic:	23.2910	-66.5332	290.6543	20.4785
*Agulhas:	23.2910	-66.5332	335.6543	65.4785
*IndianOcean:	30.3222	-59.5019	31.9043	121.7285
*NorthWestPacific:	68.9941	-20.8300	115.2246	205.0488
*NorthEastPacific:	68.9941	-20.8300	171.4746	261.2988
*SouthWestPacific:	23.2910	-66.5332	111.0058	200.8300
*SouthEastPacific:	23.2910	-66.5332	201.0058	290.8301

JPL THE CHALLENGE

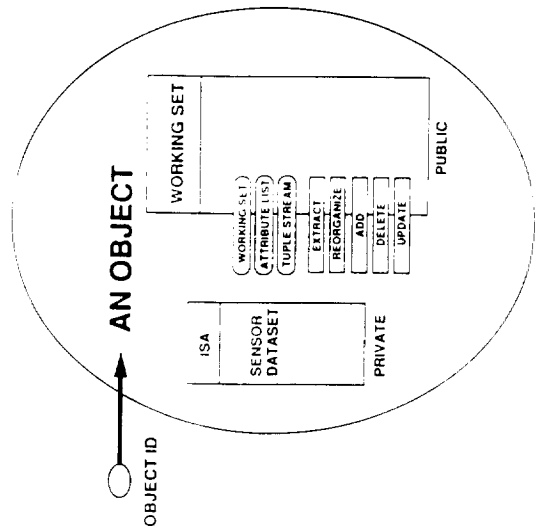
• HOW TO SIMULTANEOUSLY:

- MANAGE DISTRIBUTED, HETEROGENEOUS, INDEPENDENT DOMAINS AND RESOURCES
- PROVIDE USERS WITH TOOLS THAT ARE RESPONSIVE, HIGH-PERFORMANCE AND EASY TO USE



SERVER - PROCESS SERVER OR DATABASE SERVER

JPL DATA TRANSLATION AND REFORMATTING



S23-52

E

P. 11

DATAHUB: SCIENCE DATA MANAGEMENT IN SUPPORT
OF INTERACTIVE EXPLORATORY ANALYSIS

Thomas H. Handley, Jr. *

Mark R. Rubin **

Jet Propulsion Laboratory
Pasadena, California 91109Setting the Stage - The IssuesAbstract

The DataHub addresses four areas of significant need: scientific visualization and analysis; science data management; interactions in a distributed, heterogeneous environment; and knowledge-based assistance for these functions. The fundamental innovation embedded within the DataHub is the integration of three technologies, viz. knowledge-based expert systems, science visualization, and science data management. This integration is based on a concept called the DataHub. With the DataHub concept, science investigators are able to apply a more complete solution to all nodes of a distributed system. Both computational nodes and interactive nodes are able to effectively and efficiently use the data services (access, retrieval, update, etc.) in a distributed, interdisciplinary information system in a uniform and standard way. This allows the science investigators to concentrate on their scientific endeavors, rather than to involve themselves in the intricate technical details of the systems and tools required to accomplish their work. Thus, science investigators need not be programmers. The emphasis is on the definition and prototyping of system elements with sufficient detail to enable data analysis and interpretation leading to information. The DataHub includes all the required end-to-end components and interfaces to demonstrate the complete concept.

It is difficult, if not impossible, to apply existing tools for visualization and analysis to archived science instrument data [2]. This difficulty is generally the result of (1) incompatible data formats and the lack of available data filters; (2) the lack of true integration between the visualization and analysis tools and the data archive system(s); (3) incompatible and/or non-existent metadata; and (4) the exposure of the scientist to the complexities of networking. These problems will be multiplied by the avalanche of data from future NASA missions [8, 32]. New modes of research and new tools are required to handle the massive amount of diverse data that are to be stored, organized, accessed, distributed, visualized, and analyzed in this decade [4, 26].

The areas of most immediate need are: (1) science data management; (2) scientific visualization and analysis; (3) interactions in a distributed, heterogeneous environment; and (4) knowledge-based assistance for these functions. The fundamental innovation required is the integration of three automation technologies: viz. knowledge-based expert systems, science visualization, and science data management. This integration is based on a concept called the DataHub.

With the DataHub, investigators are able to apply a complete solution to all nodes of a distributed system. Both computational nodes and interactive nodes are able to effectively and efficiently use the data services (access, retrieval, update, etc.) in a distributed, inter-disciplinary information system in a uniform and standard way. This enables the investigators to concentrate on their scientific endeavors, rather than to involve themselves in the intricate technical details of the systems and tools required to accomplish their work; thus, investigators need not be programmers.

DataHub addresses data-driven analysis, data transformations among formats, data semantics preservation and derivation, and capture of analysis-related knowledge about the data. Expert systems will provide intelligent assistant system(s) with some knowledge of data management and analysis built

* Technical Group Supervisor
Member of AIAA

** Member of Technical Staff

Copyright © 1993 by the American Institute of Aeronautics and Astronautics, Inc. No copy right is asserted in the United States under Title 17, U.S. Code. The U.S. Government has a royalty-free license to exercise all rights under the copyright claimed herein for government purposes. All other rights are reserved by the copyright owner.

in. Eventually DataHub will incorporate mature expert system technology to aid exploratory data analysis, i.e., neural nets or classification systems. Additionally, as a long term goal, DataHub will be capable of capturing and encoding of knowledge about the data and their associated processes. The DataHub provides data management services to exploratory data analysis applications, i.e., LinkWinds [23], PolyPaint+ [15], exploratory data analysis environments.

In developing DataHub we utilize the problems as posed by the science co-investigators to aid in directing capability and development decisions. DataHub's general problem-solving structure will be applied in the general science problems, as described by the science co-investigators.

Goals and Objectives

Our goal is to integrate the results from science data management, visualization, and knowledge-based assistants into a scientific environment; to demonstrate this environment using real-world NASA scientific problems; and to transfer the results to science investigators in the appropriate disciplines.

The specific objectives of the DataHub work are to:

1. Define and develop an integrated system that is responsive to the science co-investigator's needs.
2. Demonstrate the interim capabilities to the participating science users of the system in order to receive their suggestions.
3. Transfer the results of this effort to a broad base of science investigators as appropriate.
4. Provide a system that will enable the science investigator to obtain publishable scientific information.

Emerging Relationships

As illustrated in Figure 1, LinkWinds is providing two functions: (1) a visual data exploration or analysis environment; and (2) visual browsing and subsetting services. In the first function, LinkWinds will be notified via a message of the presence of data. The existence of this data will be incorporated into the LinkWinds database menu and, hence, be made available to the user immediately. The second function will be used when it is more convenient to graphically select the subsetting attributes. After selection of the attributes, a message will be sent to DataHub, the filtering accomplished, and the results re-submitted to LinkWinds for analysis.

A new link is being established with PolyPaint+. PolyPaint+ will provide a interactive visualization of

complex data structures within three-dimensional data fields, in addition to visual subsetting services. Interactions with PolyPaint+ will require DataHub to expand its understanding of formats and data, and to provide different filtering capabilities.

The application of machine learning techniques to feature recognition in datasets of interest at JPL. The specific problem is to detect and categorize small volcanoes on Venus using the Magellan SAR data. The techniques is user interaction for feature selection and machine learning will be directly applied to the pre-processing tools used in the DataHub environment.

The Navigation Ancillary Information Facility provides a capability called SPICE (Spacecraft, Planet, Instrument, C-matrix, and Events)[19]. SPICE contains all the ancillary data associated with a mission. The data along with an extensive library are available concerning an expanding set of missions. The SPICE capability, initially developed to support science analysis, is now available as a toolkit. It is our intention to investigate the use of the SPICE toolkit in association with other applications to provide needed ancillary data and processing.

Approach

We have analyzed the management of distributed data across different computing and display resources. Subsequent to this analysis and design, we implemented the specific components required to provide needed science functions. Several prototypes have been provided to illustrate the capabilities. Additionally, we have attempted to apply knowledge-based expert system and machine learning technologies to provide "assistants" for the science investigator in data discovery and selection, tools selection and science processing. Today's solution, DataHub, takes the first steps toward the integrated solution needed to provide the means to satisfy the technology and science requirements in the 1990s by providing a high performance, interactive science workstation with the capabilities to handle both exploratory data analysis and science data management.

The Basic Concept

Figure 1 depicts the current functional architecture for the DataHub. The major functions of the DataHub include providing (1) an interactive user interface; (2) a command-based query interface; (3) a set of data manipulation methods; (4) a metadata manager; and (5) an underlying science data model. The interactive user interface, basic data operators and a data interchange

interface with LinkWinds have been implemented in the initial prototypes.

The command-based query interface, such as with LinkWinds illustrated by the double-headed arrow in Figure 1, is designed for the data visualization system to issue data management commands to the DataHub. The data manipulation methods provide the selection, subsetting, conversion, transformation, and updates for science data. The metadata manager captures the necessary knowledge about science data. Finally, the science data model supports the underlying object-oriented representation and access methods.

Figure 2 depicts the current software architecture. A layered architecture has been adopted for the implementation, which implies that any layer can be changed and/or replaced without affecting other layers. The top layer is the external interface that links to the human users via an interactive interface provided by DataHub or the visualization system via a connection interface. The data model is implemented in the intelligent data management layer. The data interface layer provides the physical data access functions.

DataHub Version 0.5 has been implemented and tested in the Sun SPARCstation and the Silicon Graphics environments. The implementation uses the software structures illustrated in Figure 2.

From a user's stand point, DataHub recognizes/understands several common datasets either by name or format, plus several other popular formats. The datasets include MCSST, CZCS, Voyager, Magellan, AVIRIS, Viking, and AirSar; the formats include VICAR [17], DSP, HDF [24], netCDF [20, 27], and CDF [3]. Present preprocessing capabilities are data filters, e.g., temporal or band selections, subsampling and averaging options, and spatial subsetting. With the data link with LinkWinds, the user may select and process a dataset of interest then proceed to the LinkWinds environment for exploratory data analysis.

The current DataHub user interface and a typical user session including interactions with LinkWinds are illustrated in Figures 3 and 4 respectively. A description of the interface design update and development may be found in [12].

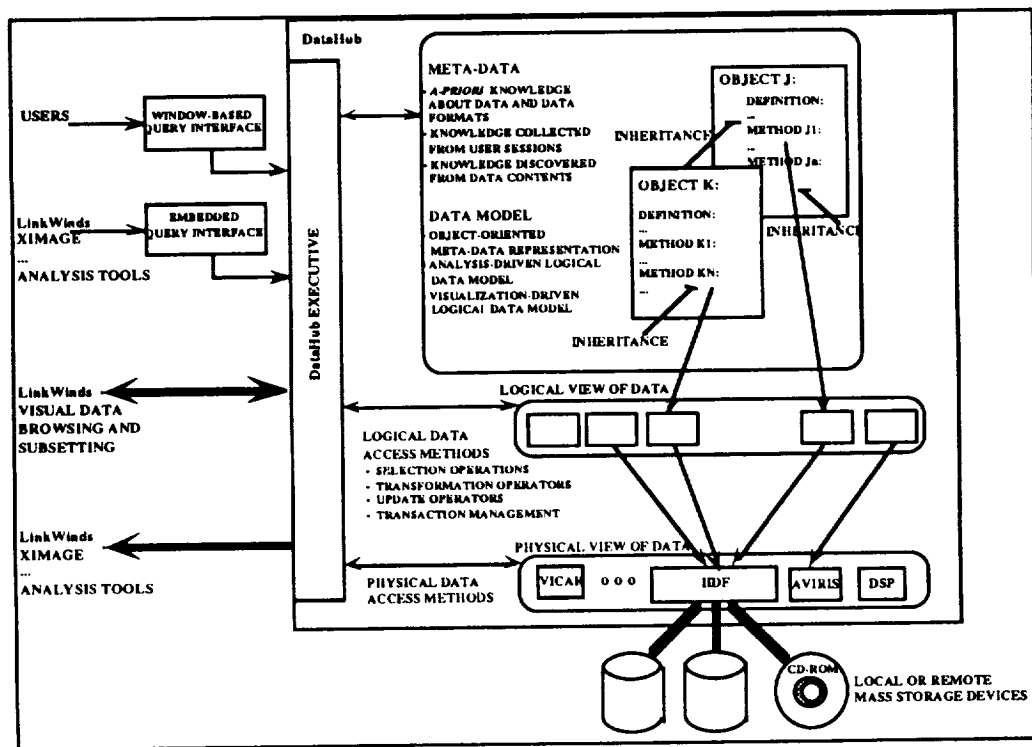


Figure 1 -- Functional Architecture

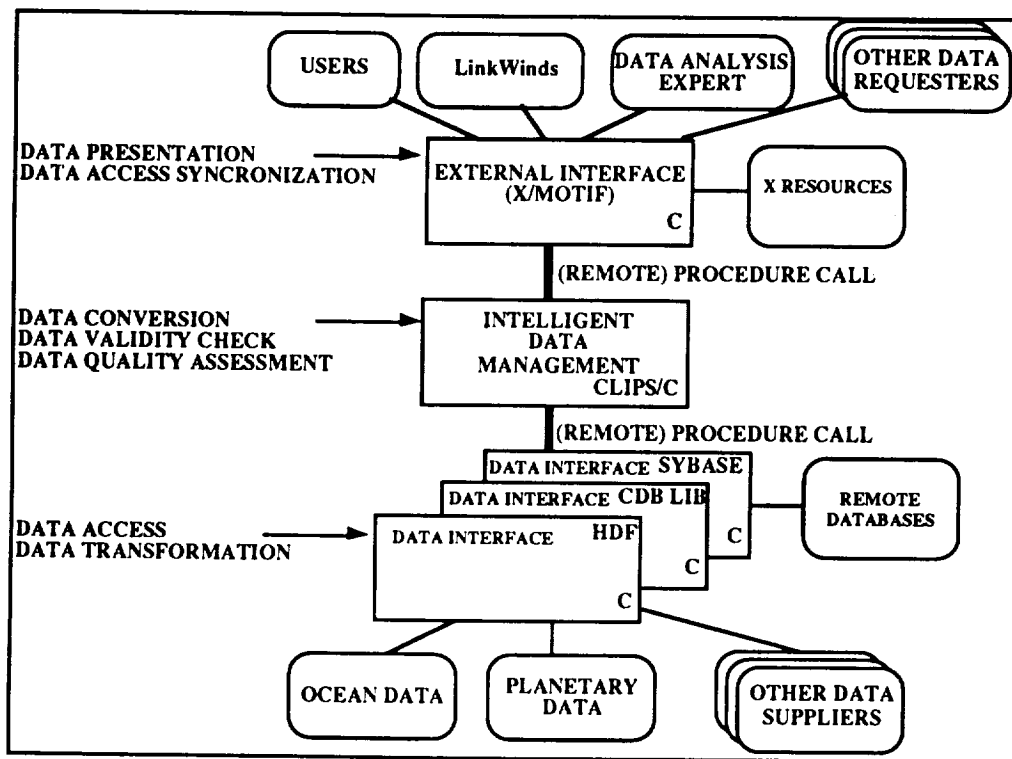


Figure 2 -- Software Architecture

Our initial experience with knowledge-based or machine learning technology was based on work accomplished using artificial neural nets. This work was spurred by our science co-investigators' needs to model regions of the ocean for which the visible and infrared imagery is obscured by clouds, and thus extrapolating biological and physical variables from cloud-free regions in space and time to the cloud-obscured regions. This produced acceptable science products but required too much technical expertise to translate into a generic tool. As described above, new machine learning techniques are being investigated to provide feature recognition capabilities with a more user-friendly interface.

A Recent Developments

Context Sensitive Help

The DataHub user interface is intended to be self-explanatory and intuitively usable with little or no instruction. In the area of user interfaces, however, intent and reality often diverge.

In packaging DataHub for distribution to a user site outside the development environment, it was obvious that a traditional "README" file was needed to detail installation instructions. It was also clear that although the DataHub user interface had largely succeeded in

achieving its goal of intuitive usability, there remained a need for a small amount of instruction to get the first-time user started. While writing a short (< 10 paragraphs) explanatory document, it became obvious that this text could be integrated into the main help system that had been designed into DataHub.

A benefit of using the X Windows resource manager to control an application's user interface is the ease with which all aspects of the interface can be customized. Textual material can be modified as simply as more traditional customizable user interface elements such as colors and layout. Because of this, any instructional text that might otherwise be included in a separate help document (either hard-copy or on-line) can be easily integrated as a dynamic part of the application itself, and eliminate the problems of help being unavailable or not findable when needed.

At the same time, a full-blown hypertext system is neither needed or appropriate for DataHub. Help for DataHub falls into two categories: Initial, new user help, and context-based help for particular DataHub capabilities. The former can be satisfied by a fairly large (as dynamic, on-screen help texts go) set of instructions, and the latter by small explanations easily accessible while the user is performing, or contemplating performing, a DataHub operation. In particular, the navigation of a help system is replaced by the navigation

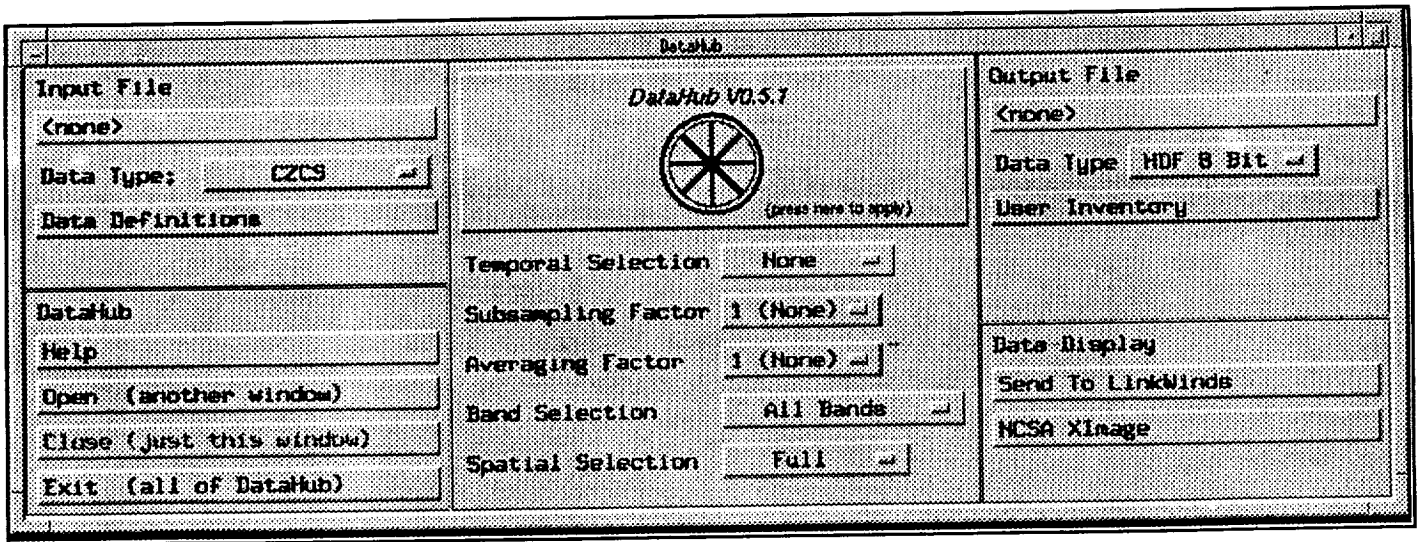


Figure 3 -- Current DataHub Interface

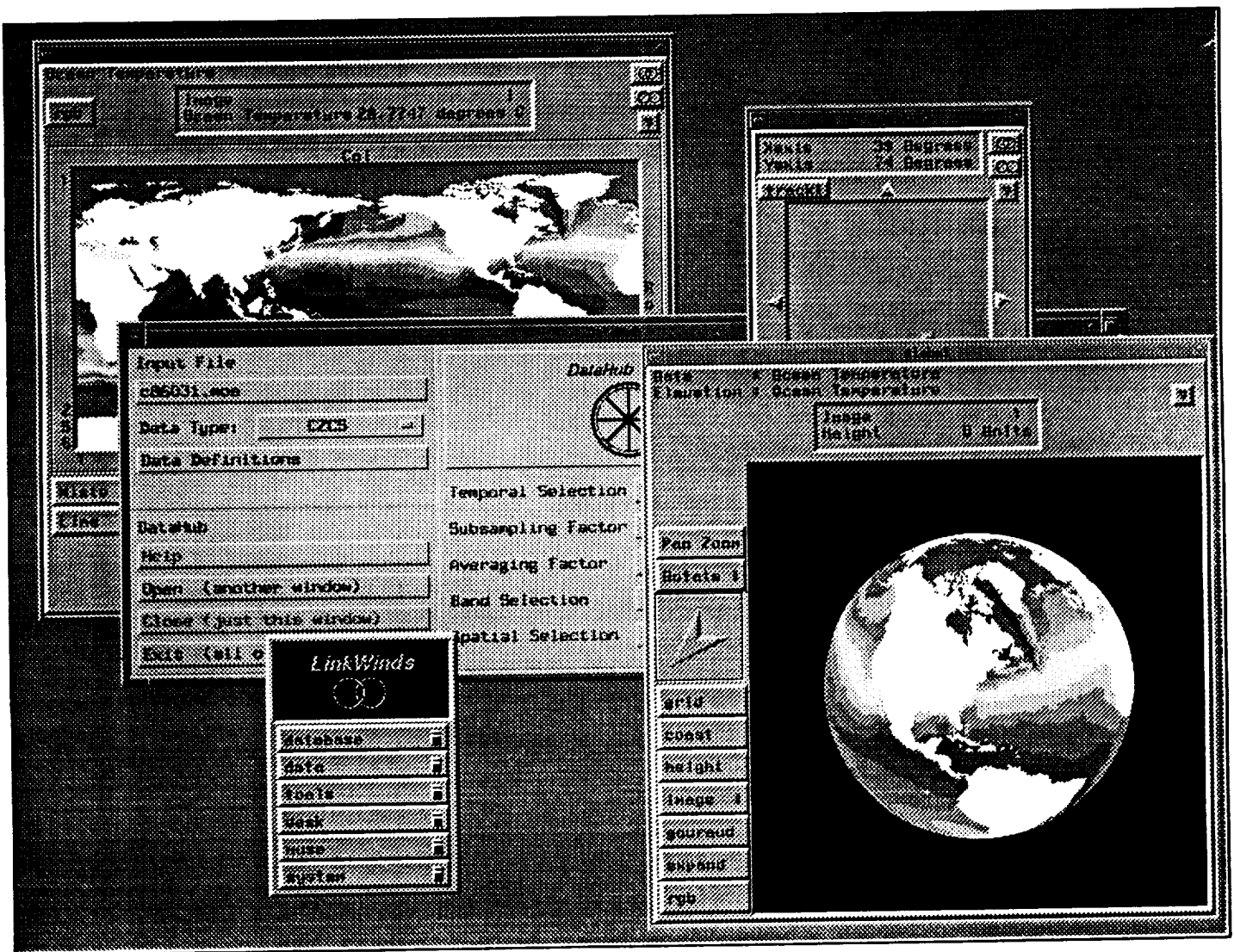


Figure 4 - Typical User Interaction

of the DataHub user interface itself, with single-level help available at each node of the interface.

Multiple, individual, help buttons fit naturally in many parts of the DataHub user interface. A help pulldown menu was added to the section of main DataHub window devoted to generic DataHub control issues. It is in this menu that an item for popping up the introductory text was placed. Additionally, all normal DataHub popup windows have help buttons that popup text dialogs containing help on their particular subject.

More difficult was deciding how to access help for graphical user interface elements (i.e. for the interface element's operations) in cases where the interface was a single button or menu with no place for a separate help button. Pulldown menus can have an additional help item added; simple pushbuttons cannot.

A context help mechanism was implemented for the case of pushbuttons, see Figure 5. The user selects "Context Help ..." from the main help pulldown menu. DataHub acknowledges this input by changing the mouse cursor to a question mark ("?") shape.

The user can then move the question mark cursor to any element of the DataHub interface, and release it to see a help dialog about that element. The underlying code sends a message requesting help to the object representing the graphical element, which in turn displays its textual help.

This method handles any and all kinds of graphic elements, regardless of their screen real-estate limitations. In fact, in the case where an element has a dedicated help button, the context-help method also works, invoking the same message and displaying the same help dialog.

Additionally, help hierarchies are a natural by-product of this implementation. Dropping the question mark cursor onto a graphical element gets help on that subject. Dropping it into the area surrounding the element gets more generic help on the type of interaction the element is a part of. For example, selecting "Subsampling Factor" or "Averaging Factor" displays help on their respective topics, but selecting physically between the two displays help on the subject of subsetting data in general.

The help system can grow and evolve using this framework. If the user drops the question mark cursor onto a graphical element that does not have a help message defined, the message automatically propagates to the ancestor of the element, repeating this process if necessary until it finds one that does have a defined help method. In this way, the user can get help (although

more general) even when specific help is yet to be implemented.

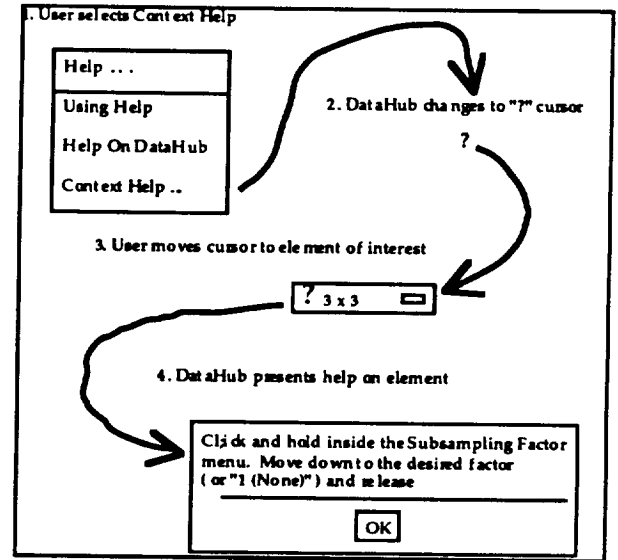


Figure 5 Context Sensitive Help

Portability

Since the goal is to provide an extensible system capable of evolving to provide solutions to broader science and engineering domains, portability is a significant issue. Initially, we conceived using a combination of C, PROLOG, and Common Lisp for the implementation. Today, portability and minimizing the cost to the user is being addressed by using common platforms (viz.. SUN SPARC stations, Silicon Graphics) and portable and public domain tools (viz. C/C++, FORTRAN, X/MOTIF, CLIPS, UNIX and SQL database interface).

netCDF Data Format

The data format Network Common Data Form (netCDF) was developed by the Unidata Program, sponsored by the Division of Atmospheric Sciences of the National Science Foundation. The emerging standard is distributed as an I/O library which stores and retrieves scientific data structures in self-describing, machine independent files. DataHub now recognizes this format.

The current implementation supports

- recognition of netCDF as a file type.
- a set of rules for conversion of netCDF to and from HDF format.

This new capability has been included to facilitate the use of netCDF data in LinkWinds and HDF data in the PolyPaint+ environments.

At this time, netCDF can be seen as providing richer structures. This is supported by the breadth of metadata annotations available as native functions. We found translation from HDF to netCDF more straight forward than the reverse.

What needs to be done?

From the design point-of-view, we have defined a general framework for science data management, and identified a critical subset of data operators for the science data applications. From an implementation perspective, we have developed prototypes that enable validation of basic concepts of data resource sharing between the data suppliers and data consumers (e.g. a data visualization system such as LinkWinds).

Based on the object-oriented design of DataHub, it is straight forward to extend the data model to capture the definitions of an existing relational data system. For example, the comprehensive data catalog built by the Planetary Data System (PDS) will become part of DataHub's data model with specialized data access methods defined to access the existing information in PDS via a standard SQL interface. This approach makes discipline-oriented knowledge readily available to DataHub. Additionally, expanded knowledge about data formats and data semantics in various science disciplines will be built into DataHub. It is a goal that the understanding of the visualization and analysis tools will also become part of DataHub such that special data operators will be built automatically using basic known operators. The data quality assessment issue of science data after data transformation will be a research area for DataHub, and will be addressed in the next steps.

We will enhance the existing prototype to provide access to additional data sets while expanding the capabilities for direct support to the science co-investigator. Particularly, the issues associated with processing multi-spectral data will be addressed. We will be enhancing the preprocessing capabilities by accessing and utilizing the NAIF SPICE ancillary data as it become available.

Besides continuing to evolve to a more object-oriented implementation, several issues will be addressed. When data transformation or conversions take place, we need to assure the preservation of data validity or quality measures. We need to treat the data quality assessment issues such as (1) treatment of missing data and (2) data quality associated with data interpolation, data transformation, etc.

Expanded knowledge about the data is of significant importance. This includes knowledge of data formats (e.g. usage of metadata embedded in the data set

headers), data semantics (e.g. meaning of data values, relationships between data sets, discipline-dependent data access/analysis methods) and data semantics as represented by the users' context in the visualization regime (e.g. what are the links, dataflows, etc. as encapsulated in the LinkWinds environment). The ability to detect and understand this expanded knowledge will be incorporated into the label-understanding expert system.

Additional understanding of the analytical tools required for data selection, data transformation and data conversion in order to support the visualization requirements is needed. These may be thought of as filtering tools to select and prepare data for use in the visualization environment. These additional tools will be defined and implemented.

In those cases where selection criteria are so complex that they are most easily exercised visually, it is clear that a close integration of the database management system, and the data visualization system is advantageous. Such integration will be studied by closely tying DataHub with LinkWinds so that DataHub will be accessible from LinkWinds and LinkWinds will be accessible from DataHub, each being used to best advantage in the data management processes.

Finally, we will address the issues associated with data presentation. In particular, data exchange protocols that facilitate visualization are to be addressed first.

Major Components

DataHub will be enhanced to include these capabilities:

- interactions to support finding, selecting and processing multi-spectral datasets (initially AVIRIS).
- band aggregations
- band filters (e.g., removal of artifacts of the instrument)
- 3D subsetting/averaging
- journal and transaction management will playback capability.
- expanded data model that includes user-defined data conversions.
- canonical set of data objects and methods
- self-describing data objects and methods
- user defined defaults for spatial regions, temporal periods, etc.
- incorporate the metadata into the interfaces with LinkWinds and PolyPaint+.
- expanded rule-based capability to understand foreign datasets, leading to a capability for interpretative conversions and transformations.
- expanded data dictionary for use in label recognition, plus the ability to dynamically add

new object attributes once their semantics are clearly understood.

- initial usage of calibration and registration data
- quality measures, to include
- processing lineage
- null and missing value recognition and usage in processing
- incorporation of content-based applications such as the machine learning capability described above
- expanded interactions with LinkWinds and PolyPaint+
- distribution of DataHub processing and interactions and remote services.

Using the DataHub, scientists will request data for presentation and analysis in a specific way for use in their applications, without being particularly concerned with the original location and format of data being utilized. Applications adhering to the DataHub protocols and interfaces may interoperate sharing results through the DataHub.

As described previously, LinkWinds will be enhanced to have two-way communications with DataHub. Besides receiving the user's selected data for analysis, LinkWinds will provide graphical subsetting and transformation parameters and send a processing request for DataHub to execute and return the desired data.

PolyPaint+ will have a similar interface as LinkWinds. After this communications and processing link has been implemented, DataHub will be enhanced to provide more specialized processing for the PolyPaint+ community (that is say, netCDF, super computing, and modeling).

Machine Learning and Feature Detection

It is difficult for a scientist to examine and understand data with a large number of dimensions. Scientific visualization tools are one means for performing necessary transformations and dimensionality reduction to allow a scientist to "see" meaningful patterns in the data. However, these require that the scientist specify the necessary steps. Faced with multi-spectral remote-sensing data arriving over more than 200 channels, expecting a scientist to study the entire data set becomes unreasonable. This often results in using only parts of the data channels or using the data in very limited ways. An automated tool for aiding the analysis of such high dimensional data sets would enable scientist to get at more of the information contained in the data.

We will use machine learning and pattern recognition techniques to aid in the analysis of multispectral data. Consider a scientist interested in characterizing certain

regions in the data, for example, locating the areas on earth where certain minerals are present, or where some phenomenon of interest occurred. By selecting portions of the data of interest and others that do not contain phenomena of interest, a scientist is essentially pointing out examples (instances) of the desired target. These can be treated as training data, and used by learning algorithms to automatically formulate classifiers that can detect other occurrences of the target pattern in a large data set. Furthermore, since the learning algorithms are capable of examining a large number of dimensions at once, they may be able to find patterns that would be too difficult for a scientist to derive by manual analysis. In a sense, this offers the option for a "logical" versus a "visual" visualization of the patterns in the data. That is, the algorithms produce a characterization of subsets of interest in the data in terms of logical expressions involving multiple input variables (channels). Often, it is possible to express such patterns in terms of compact rules involving an unexpectedly small number of variables. For example, channels 104 and 202 being in certain ranges may be highly predictive of a phenomenon that the scientists could not easily characterize using the first six channels.

The use of learning algorithms thus provides flexibility in terms of adapting to a wide variety of detection problems. Our decision tree based learning algorithms produce rules that are easily examined and understood by humans. This contrasts with a statistical regression or neural network based approach, where the resulting forms are difficult to interpret.

Distributed Blackboard System

The blackboard model allows for a flexible architecture with diverse knowledge sources cooperating to formulate a solution opportunistically [16]. A distributed blackboard system running across multiple workstations can allow multiple scientists in different physical locations to work together on a single problem cooperatively.

The DataHub metadata manager has been ported to a distributed environment across multiple Sun SPARCstations [22]. This distributed environment is the underlying layer of an ongoing distributed blackboard implementation. It is expected that the DataHub system can sit on top of this blackboard system to function as a Groupware for multiple scientists from multiple science disciplines.

With this capability, DataHub can distribute the data access and data conversion load across multiple computers. At the same time, multiple users can access multiple data sources via this distributed scheme of DataHub.

With the distributed blackboard, DataHub can have multiple data servers with metadata (i.e., discipline knowledge) about multiple data sources sitting across the network. Each data server acts as an independent knowledge source in the blackboard system. The DataHub data servers use a consistent data access mechanism provided by DataHub. The scientists use a consistent user interface of DataHub even though they are running the DataHub data client on their own workstations geographically separated from one another.

The inter-disciplinary knowledge about data can be stored in higher level knowledge sources (i.e., agents) in the blackboard system. Whenever a scientist has a need of a dataset that is outside a single discipline, this inter-disciplinary knowledge source is utilized to provide intelligent data access capability to access the right data from the right source.

The distributed blackboard is implemented using a reliable distributed computing protocol provided by Cornell's ISIS [1, 7]. ISIS version 2.1 is in the public domain. The concept of having process groups in a distributed environment with guaranty on message arrival sequence for messages from multiple senders fits the need of the blackboard implementation.

Development and Deliverables

We have planned three steps in the next phase of DataHub prototyping:

Step 1.

- Design and develop DataHub processing of multi-spectral data sets for the science co-investigator.
- Initiate the distribution of DataHub processing and provide general remote services.
- Design and develop interfaces to PolyPaint+. Collect functional requirements from the user community.
- Design and develop the machine learning interface.
- Demonstrations will use the data sets as determined by the science co-investigator.

Step 2.

- Provide data abstraction and knowledge engineering to support applications in the LinkWinds and PolyPaint+ environments.
- Demonstrations will use the data sets as determined by the PolyPaint+ user community.

Step 3.

- Provide the knowledge engineering required to utilize the computing environment and its tools. Incorporate this knowledge into the DataHub.
- Provide support within the DataHub of all the required datasets (homogeneous/regular and heterogeneous).

- Demonstrations will use the data sets as determined previously.

The development cycle used to solve the problems addressed above will be to: define/expand the science co-investigator's problem; design, implement, integrate test, demonstrate, evaluate, and transfer to the scientist co-investigator; and then iterate these steps. In each cycle these areas will be addressed: (1) The DataHub; (2) Knowledge-based assistance for the DataHub; (3) Machine learning for feature recognition; (4) A problem posed by a science co-investigator; and (5) LinkWinds/PolyPaint+ interface and protocol.

An incremental development methodology will be utilized: "do-a-little, test-a-little".

Throughout the implementation effort, the science co-investigator and other scientists will participate in the design. This feedback and evaluation is important in providing a product that contributes to the scientists' ability to accomplish their science objectives. The success of the proposed work will be measured by the science utility of the work products.

Benefits and Expected Results

The principle product of the proposed work is the demonstration of an integrated environment in which a science co-investigator will be able to accomplish data analysis and interpretation leading to publishable scientific information. Thus, DataHub is addressing broad aspects of:

1. Providing innovative ways to facilitate the scientific endeavor or "mean-time to discovery" [33] when working with large volumes of data. The traditional computing data life-cycle is typically a sequential process. This traditional view provides sequential support to what is actually a highly-interactive, iterative process. DataHub will provide a data life-cycle as illustrated in Figure 3.
2. Providing access to remote data, local data filtering and management and interactive exploratory data analysis.
3. Applying knowledge-based expert systems and machine learning at the original data selection, in intermediate data filtering and in rule-based applications.

The DataHub will provide an end-to-end solution to problems of this generic type, thus enabling science investigators to produce higher-level products through an analysis environment which provides an integration of required functions. This environment consists of:

1. An interface between the scientific visualization and analysis environment and the data required to perform the analysis.
2. Expert system / knowledge engineering-based analysis assistants and machine learning techniques to do:
 - data discovery and data selection
 - feature and image understanding preprocessing
 - visualization and analysis tool selection
3. The LinkWinds and PolyPaint+ environments and their analysis tools as the visualization mechanism and user interface environment.

The benefits to NASA deriving from the DataHub include:

1. Ability to analyze massive volumes of data in a cost-effective manner.
2. Freedom for the NASA mission scientists to do the interpretative, creative aspects of science work.
3. An advanced prototype for science support.
4. Availability of common system modules and data formats for other developers.

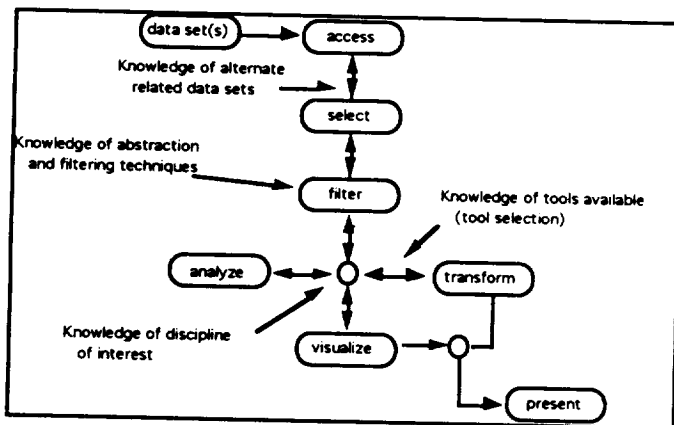


Figure 3 -- Knowledge-Based Visualization and Analysis

Acknowledgments

This work was supported by NASA through Joe Bredekamp's Applied Information Systems Research program. The authors wish to thank the scientist users of DataHub for their valuable insights into discipline-specific needs in data usage and semantics and user interactions.

The research described in this paper was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

If you have comments, questions, or would like to discuss the uses of DataHub, the authors may be reached at:

T. Handley
 thandley@spacemouse.jpl.nasa.gov
 M. Rubin
 mark@phineas.jpl.nasa.gov

Bibliography

1. K. Birman, A. Schiper, P. Stephenson, "Lightweight Causal and Atomic Group Multicase", ACM Transactions on Computer Systems, Vol. 9, No. 3, August, 1991.
2. M. Botts, Informal Communication "Results of a NASA survey for EosDIS ", University of Alabama, June 2, 1992.
3. CDF User's Guide for UNIX Systems Version 2.1, January 9, 1992, National Space Science Data Center, NASA - Goddard Space Flight Center, Greenbelt, MA.
4. Joseph C. Ciarrarano, CLIPS User's Guide Version 5.1, September 10, 1991, Volumes 1-5, Johnson Space Center, Information Systems Directorate, Software Technology Branch.
5. Peter J. Denning, "Information Technologies for Astrophysics Circa 2001- A Position Paper", NASA Workshop on Astrophysical Information Systems, May 23-25, 1990.
6. Stephanie Granger-Gallegos, et. al., "Automatic Data Distribution for a Large Geophysical Data Set", submitted to The Earth Observer, July 1992.
7. O. Hagsand, H. Herzog, K. Birman, R. Cooper, "Object-Oriented Reliable Distributed Programming", 2nd International Workshop on Object-Orientation in Operating Systems, 1992.
8. Thomas H. Handley, Jr., Larry E. Preheim, "High Rate Information Systems: Architectural Trends in Support of the Interdisciplinary Investigator", AIAA/NASA Second International Symposium on Space Information Systems, September 1990.
9. Thomas H. Handley, Jr., Larry E. Preheim, "NASA Information Systems Commonality and Convergence", AIAA Computing in Aerospace 8, October 1991.
10. Thomas H. Handley, Jr. P. Li, "DataHub: Intelligent Data Management for Data Discovery", International Space Year Conference on Earth and Space Science Information Systems, February 10-13, 1992.
11. Thomas H. Handley, Jr. Y. Philip Li, Allan S. Jacobson, An V. Tran, "DataHub: Intelligent Data Management", 1992 ASPRS/ACSM Convention, August 3 - 7, Washington, D. C.
12. Thomas H. Handley, Jr., Mark Rubin, Y. Philip Li, "DataHub - Knowledge-based Science Data

- Management for Exploratory Data Analysis", SPIE Spring 1993, Orlando, Florida.
13. Thomas H. Handley, Jr., " DataHub: Intelligent Distributed Science Data Management to Support Interactive Exploratory Analysis", Computing in Aerospace 9, San Diego, California, October 19-21, 1993.
 14. Thomas H. Handley, Jr., "DataHub: Science Data Management in Support of Interactive Exploratory Analysis", PECORA12, August 24-26, Sioux Falls, South Dakota.
 15. Klemp, J. B., M. J. McIrvin, and W. S. Boyd. 1990: PolyPaint - A Three-Dimensional Rendering Package. In Proc. Sixth International Conference on Interactive Information and Processing Systems for Meteorology, Oceanography, and Hydrology, 5-9 February 1990, Anaheim, Calif. AMS, Boston, Mass., 286-293.
 16. B. Hayes-Roth, "A Blackboard Architecture for Control", Artificial Intelligence, 26 (1985).
 17. Susan LaVoie, et. al., VICAR User's Guide Version 2, June 1, 1989, JPL D-4186, Rev A.
 18. Joel L. Morrison and Kathryn Wortman, Editors Special Issue "Implementing the Spatial Data Transfer Standard", Cartography and Geography Information Systems , Vol. 19, No. 5, December 1992.
 19. NAIF AMMOS V18.1 MO Delivery Documentation Package, Navigation Ancillary Information Facility, 25 February 1993
 20. NetCDF User's Guide Version 2.0, October 1991, Unidata Program Center, Boulder, CO.
 21. T. Handley and L. Preheim, "High Rate Information Systems: Architectural Trends in Support of the Interdisciplinary Investigator", AIAA-90-5084, AIAA Second International Symposium on Space Information Systems, 17-19 September 1990, Pasadena, CA.
 22. Y. Philip Li, "dCLIPS: A Distributed Clips Implementation", Computing in Aerospace 9, San Diego, CA, October 1993.
 23. Allan S. Jacobson and Andrew L. Berkin, "LinkWinds: A System for Interactive Scientific Data Analysis and Visualization", Proceedings of High Performance Computing Conference, Society for Computer Simulations, March, 1993.
 24. NCSA HDF Calling Interfaces and Utilities Version 3.1, July 1990, NCSA Software Tools Group, University of Illinois at Urbana-Champaign, IL.
 25. L. Preheim, T. Handley, "Data System Interoperability Task, Requirements Commonality Study Report", JPL D-7187, February, 1990.
 - 26 M. Stonebraker and J. Dozier, "Large Capacity Object Servers to Support Global Change Research", Report # 91/1, SEQUOIA 2000, University of California.
 27. Russ Rew and Glenn Davis, "NetCDF: An Interface for Scientific Data Access", IEEE Computer Graphics & Applications, July 1990, pp. 76-82.
 28. A. Silberschatz, Michael Stonebraker, Jeffery D. Ullman, "Database Systems: Achievements and Opportunities", TR-90-22 Department of Computer Sciences, The University of Texas at Austin, Report of the NSF Invitational Workshop on the Future of Database Systems Research held in Palo Alto, California, February 22-23, 1990.
 29. Spatial Data Needs: The Future of the National Mapping Program, National Academy Press, Washington, D.C., 1990
 30. Lloyd A. Treinish, "SIGGRAPH '90 Workshop Report Data Structures and Access Software for Scientific Visualization", Computer Graphics, Vol. 25, No. 2, April 1991
 31. Lloyd A. Treinish, "An interactive, discipline-independent data visualization system", Computers in Physics, Jul/Aug 1989, pp. 55-64.
 - 32 James Warner, "Visual Data Analysis into the '90s", Pixel, January/February 1990, pp. 40-44.
 33. M. Mitchell Waddrop, "Learning to Drink from a Fire Hose", Science, Vol. 248, pp. 674-675, May 11, 1990.

DataHub sponsored analysis

A. Tran

D. Collins

524-92

629-50

Introduction

The primary productivity of phytoplankton in the ocean is largely responsible for the assimilation of carbon into the oceanic environment, and thus in part the removal of carbon from the atmosphere. Because the ocean is thought to be a primary sink for atmospheric carbon, the basin wide and global distribution of oceanic primary productivity is of central importance in the global budget of carbon. To understand the global productivity of the oceans, the interactions between the physical and biological structures must be known. The biological population of the ocean is highly variable both spatially and temporally on all time and space scales. The global nature of this problem then requires the use of satellite instrumentation as the only platform capable of providing coverage on temporal and spatial scales that are appropriated to the assessment of carbon flux in the ocean. The goal of this research is to increase our understanding of the sources of variability in the sea to provide a more accurate assessment of oceanic productivity from ocean color imagery. The objectives of this research are the description of the spatial and temporal distributions and variability of the planktonic community in the sea and primary productivity of that community. To achieve these objectives, remotely sensed data of the spatial and temporal distributions of pigment concentration and sea-surface temperature are required to provide a global description of the seasonal variability of the water column primary productivity.

To address the broader context of the primary productivity of the sea, the physical and biological processes and their variability, including changes in water mass, incident irradiance, nutrients and consequent formation of blooms of difference species of marine phytoplankton and bacteria must be studied. In this investigation, we will use time series of the pigment distributions, taken from the Coastal Zone Color Scanner (CZCS), and of the sea-surface temperature, taken from the NOAA Advanced Very High Resolution Radiometer (AVHRR). These time series will be examined to determine the spatial and temporal statistics of productivity, including the interannual variations that occur in productivity caused by variations in the physical environment.

For this task we have chosen to use monthly composite global maps created from the satellite imagery. The pigment maps are created from the ratios of upwelling radiance at 440, 520 and 550 nm, and have been composited from a data set that is characterized by a sparse data coverage because of the presence of clouds and because of the sampling characteristics dictated by the Nimbus-7 satellite operations. The monthly composite images from the CZCS contain significant regions for which no data exist. Attempts to estimate the global primary productivity of the ocean from these composite images have yielded a preliminary assessment of the net annual flux of carbon from the atmosphere into the oceans to be 3.2 G-tons Carbon per year, based on estimates of the water leaving radiance, and a regression against carbon flux from the work by Mitchell, et al. 1992. To provide a better estimate, and to provide the time series of this flux, we must interpolate the pigment images to provide an estimate of the pigment concentration in regions for which the data is inadequate. Conventional techniques such as bi-linear interpolation and spline fits have given insufficient results because of the large areas of missing data.

The MCSST data product can be used to understand variations in the sea-surface temperature in regions where large data gaps are present because this product has used an interpolated data field

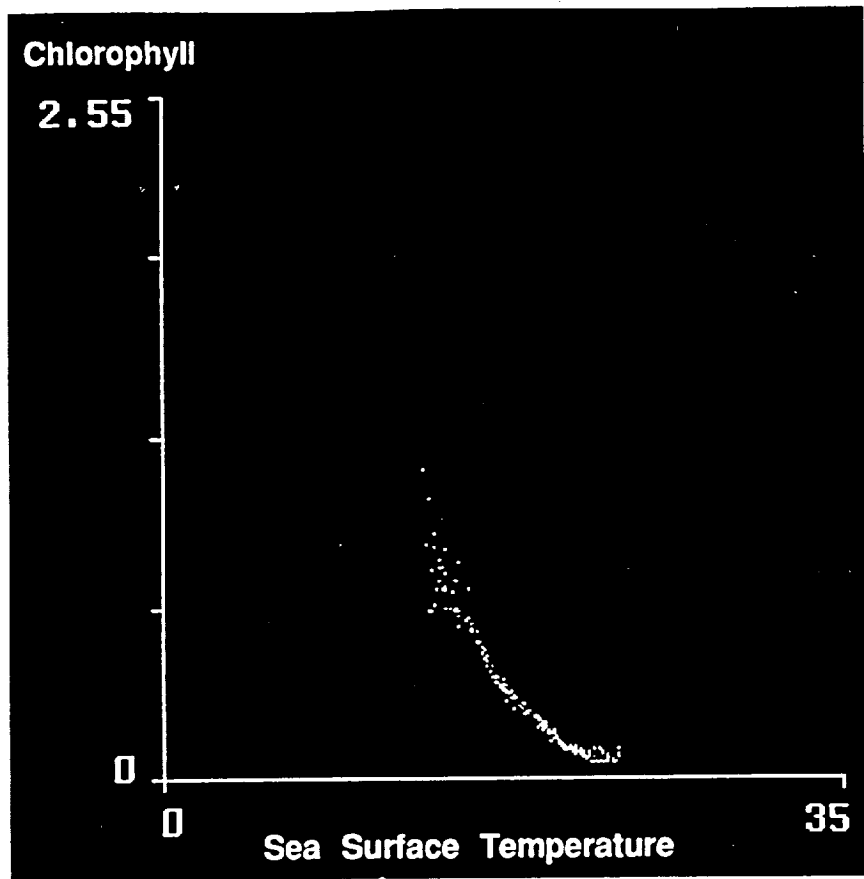


Figure 1. Correlation between chlorophyll and temperature center at 15 degrees north.

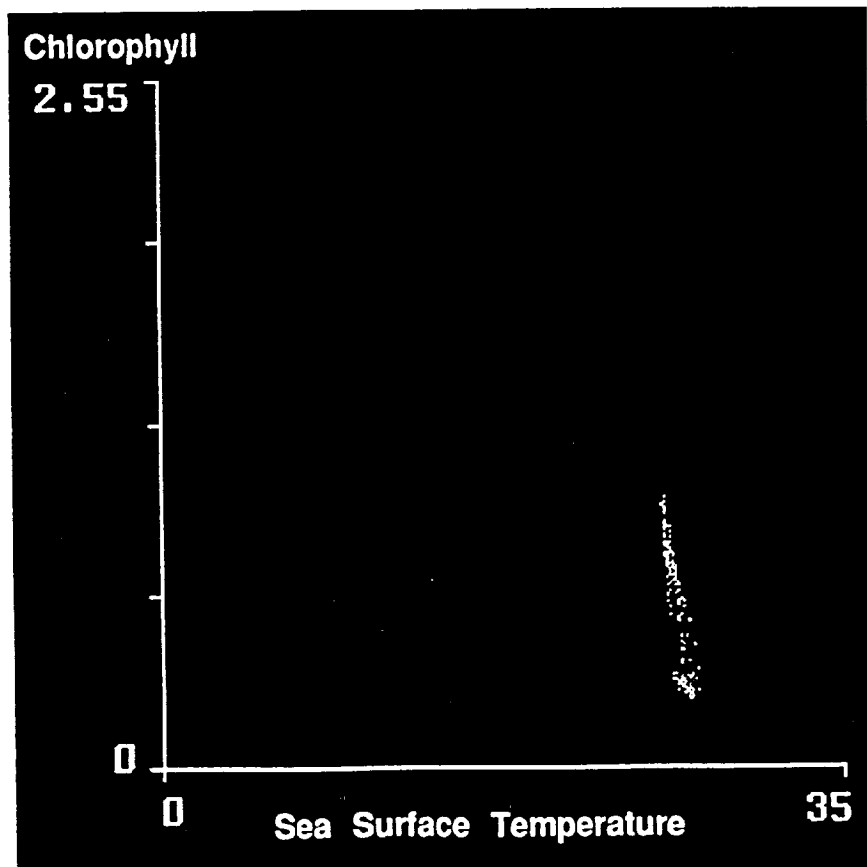


Figure 2. Correlation between chlorophyll and temperature center at 35 degrees north.

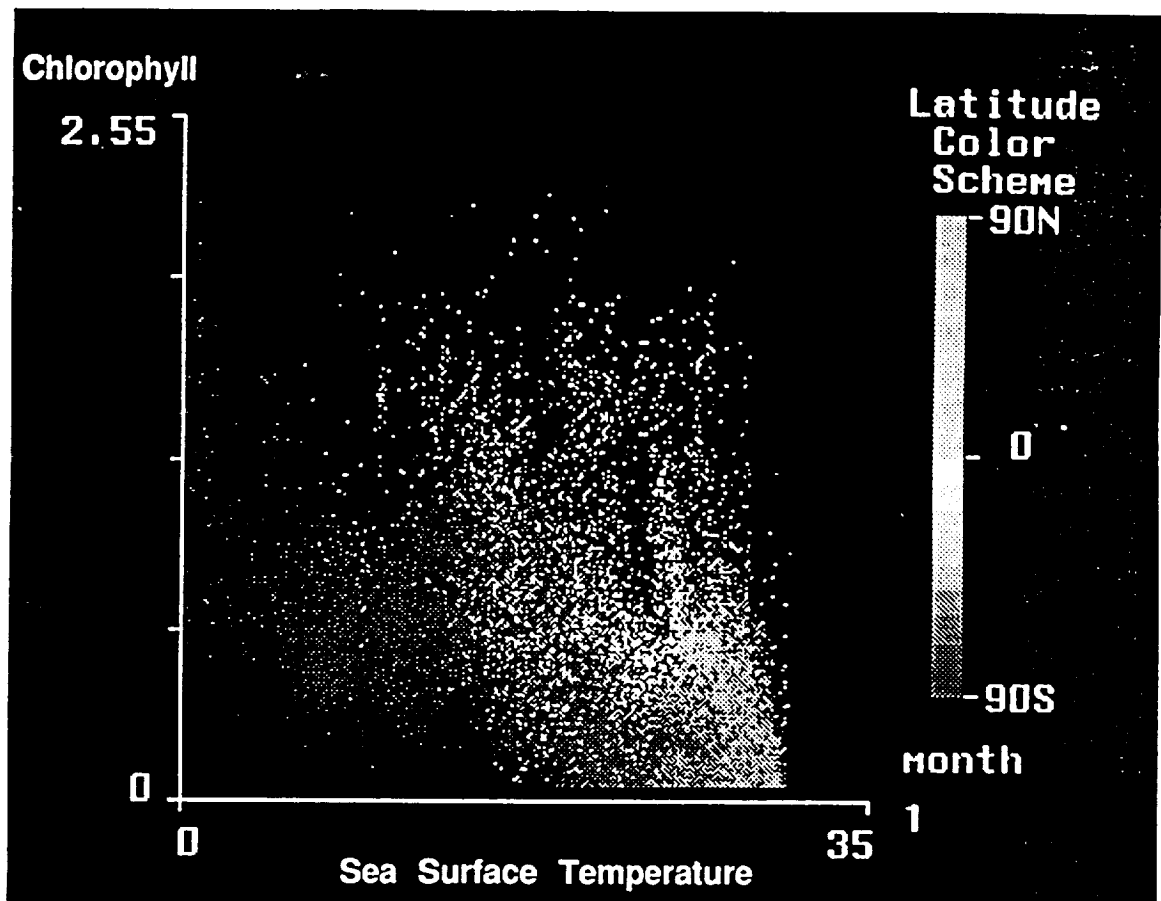


Figure 3. Correlation between chlorophyll concentration and temperature for entire global

ORIGINAL PAGE IS
OF POOR QUALITY

to supply missing data values for regions for which clouds obscure the surface during the time of the monthly composite. To provide internal consistency between the pigment and sea-surface temperature data fields, the same interpolation scheme will be used for both data fields.

The primary productivity of the sea has been shown to be related to both the standing stock of the phytoplankton population, and to the temperature of the sea, which both regulates the metabolism of the planktonic organisms and reflects the nutrient status of the sea through a physical relationship between temperature and nutrients in newly upwelled water. For these reasons, working with Mitchell, et al., we have developed a relationship between the temperature of the sea, the upwelled radiance ratio in the photosynthetic bands, and the primary productivity of the sea. While this relationship is still under investigation, the regressions that have been produced indicate that a strong correlation exists for the flux of carbon through the surface layer, and these relationships will be used to produce the first time series maps of the global flux of carbon in the sea.

Through the primary productivity, the standing stock of phytoplankton, as reflected by the pigment concentration, is also related to the temperature, although in a very complicated manner. It is this relationship that we have exploited in the interpolation of the pigment fields. Figures 1 and 2 indicate two latitudinal regions in the ocean, one at 15 degrees and one at 35 degrees north. In these figures, we illustrate two facts: First that at each latitudinal band, there is a strong correlation between temperature and pigment concentration. Second, that the correlation is very different for these two regions. The global picture for the correlation between temperature and pigment concentration is shown in Figure 3. These results indicate that we may use temperature in the interpolation of the pigment fields, but that the algorithm is both regional and seasonal in nature, leading to an exhaustive computational problem if conventional analysis were to be applied. These facts have led us to investigate different methods for the interpolation of the pigment fields in regions for which sufficient data is not available to provide a satisfactory estimate of the pigment to permit an estimate of the productivity.

The first method that we have examined is the use of a least squares regression using both temperature and pigment for the estimate. This technique will find a matrix transformation mapping spatial averages of temperature and pigment data and latitude values onto the space of pigment values such that the difference between the two sets is minimal in the root mean square sense. The variation of input parameters can be extended to include the square or cube of the spatial variables. These variables were combined in an equation where the coefficients were determined by the least squares technique. This analysis was conducted using the IMSL (International Mathematical Statistical Library) software.

Several polynomials with different variable combinations were used to examine the variability of error produced with each equation. The coefficients of these polynomials describe the contribution of each variable to the predicted pigment value. The results of studies conducted on a restricted data set indicate that a simple linear regression based on the pigment alone gives a satisfactory fit to the data from the trial cases. The results suggest further testing on a significantly larger data set, using multiple iterations of the interpolation process.

The second method uses the methods of a neural-net, coupled with a bi-linear interpolation, using

both the temperature and pigment fields to form the estimate of the missing data in the pigment field, and the temperature field alone for the estimate of the missing temperature values. This technique relies on the pigment fields both past and future for the pixel in question, the past and future temperature fields, coupled with a spatial interpolation of the pigment field to produce an estimate for the missing data pixel. The neural-net system is trained on a data set which has both the spatial and temporal coverage appropriate to the data set under investigation, and is used on the global data set for all time. The initial data set is shown in Figure 4, which illustrates the large areas of missing data in the global pigment images. The results of the interpolation are shown in Figure 5, which illustrates the degree to which the fields may be interpolated using this technique. The technique has been verified by removing data from the original data set, applying the technique to regenerate the data, and comparing the original data to that replaced by the artificial intelligence system. Figure 6 describes the correlation between the predicted pigment concentration from the Neural-Net and the pigment concentration from the satellite measurements. The correlation coefficient for this estimate is $R^2 = 0.952$.

For this task we used the most well known neural-net classifier (known as "back-propagation"). Back-propagation was introduced originally (Rumelhard, 1986), it was proposed that the criterion function to optimized using gradient descent. However, it was soon realized that more efficient algorithm and training techniques can be employed; the "momentum" term (Rumelhard, 1986) is the most popular example of such improvements. This techniques incorporated gradient descent and the previous weight change is used to update the weight vector.

We used seven different variables as input and one output into neural-net program. The inputs consist of three from the CZCS pigment field, three from the AVHRR temperature field, and the latitude of the center pixel. Figure 7 shows the lay-out input parameters for back-propagation Neural-net.

The relationship between the global chlorophyll data and the temperature product is not well defined. Neural-nets have shown the ability to handle multi-dimensional data sets with non-linear relationships.

From these experiments, we conclude that the neural net permits the computation of a globally interpolated data field for all time. The results of this study are being evaluated to determine the scientific validity of both techniques.

References

D.E. Rumelhart, G. E. Hinton, and R. J. Williams, 1986: "Learning internal representations by error propagation" in Parallel Distributed Processing. MIT press, Cambridge MA, Chapter 8, 318-362, 1986.

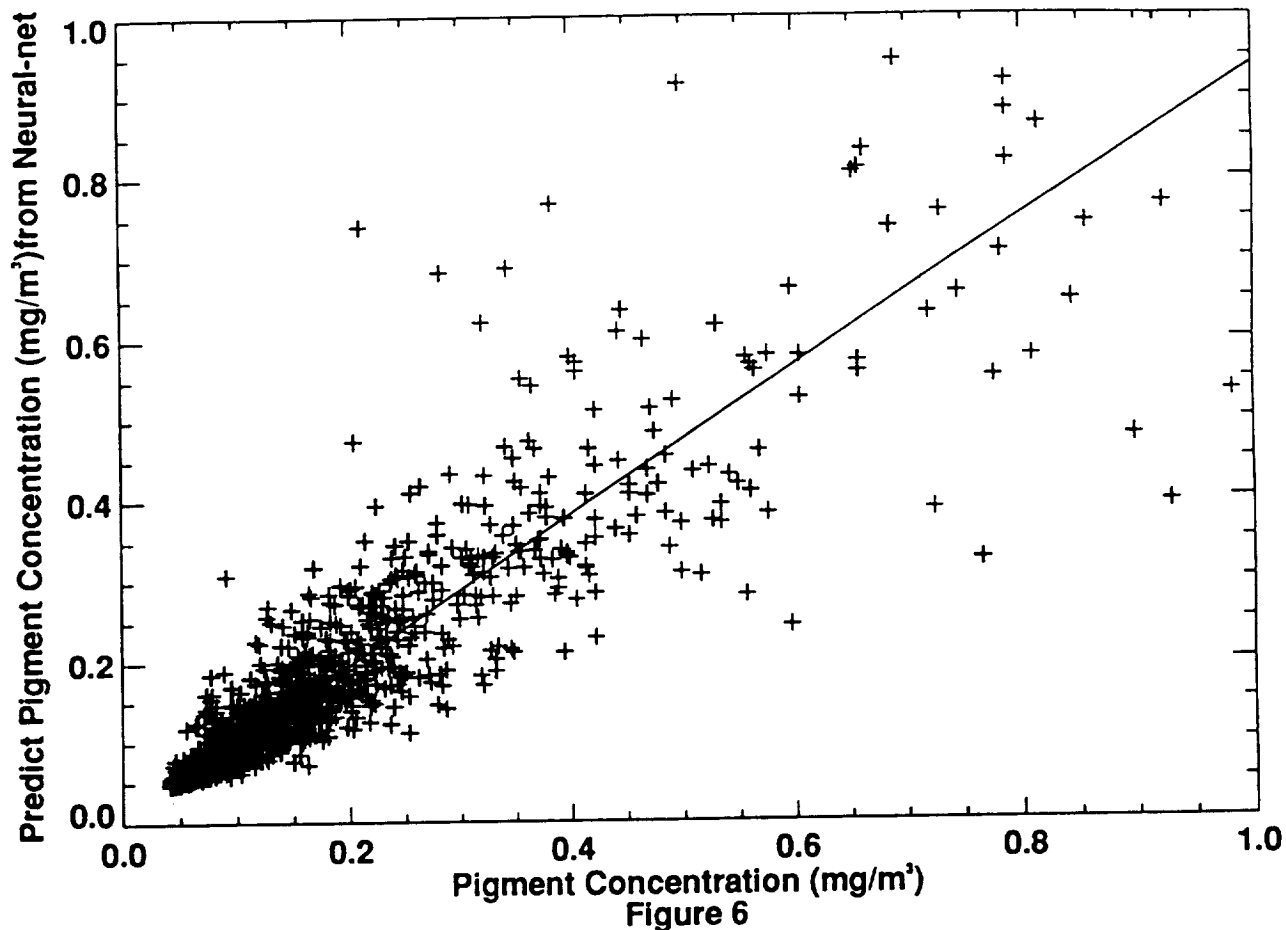


Figure 6

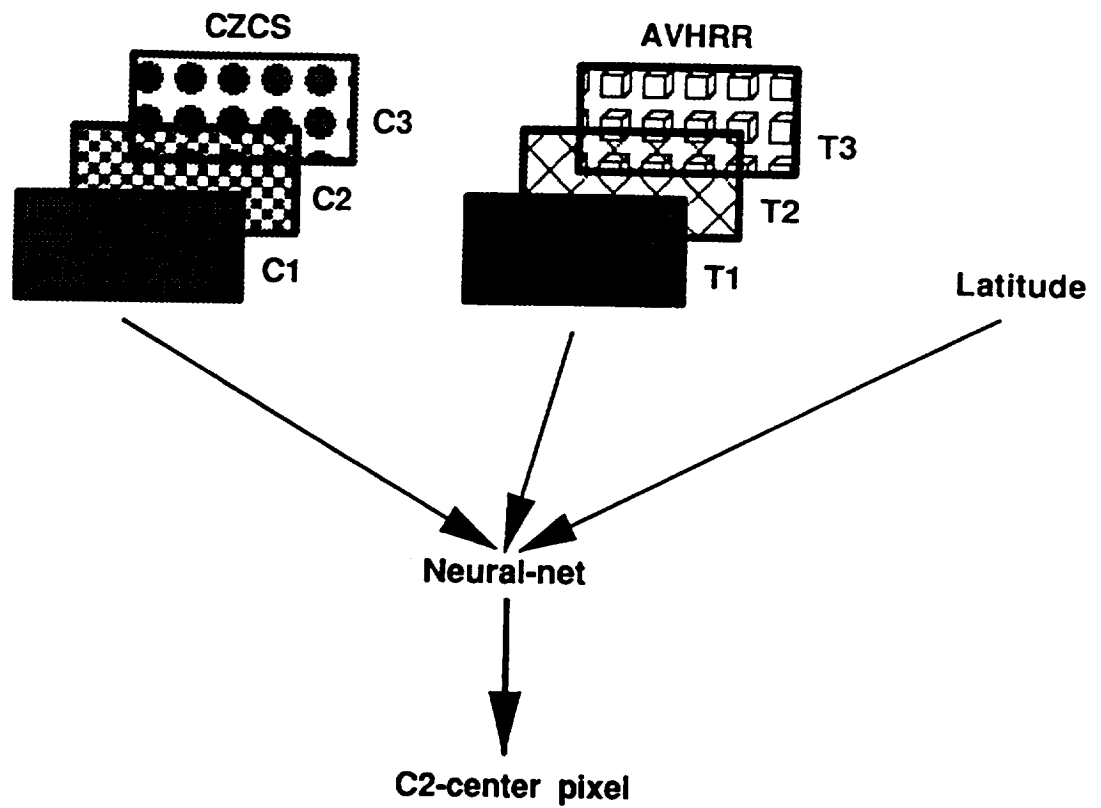
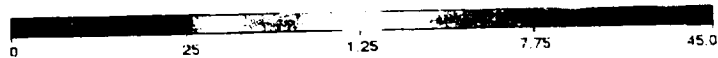
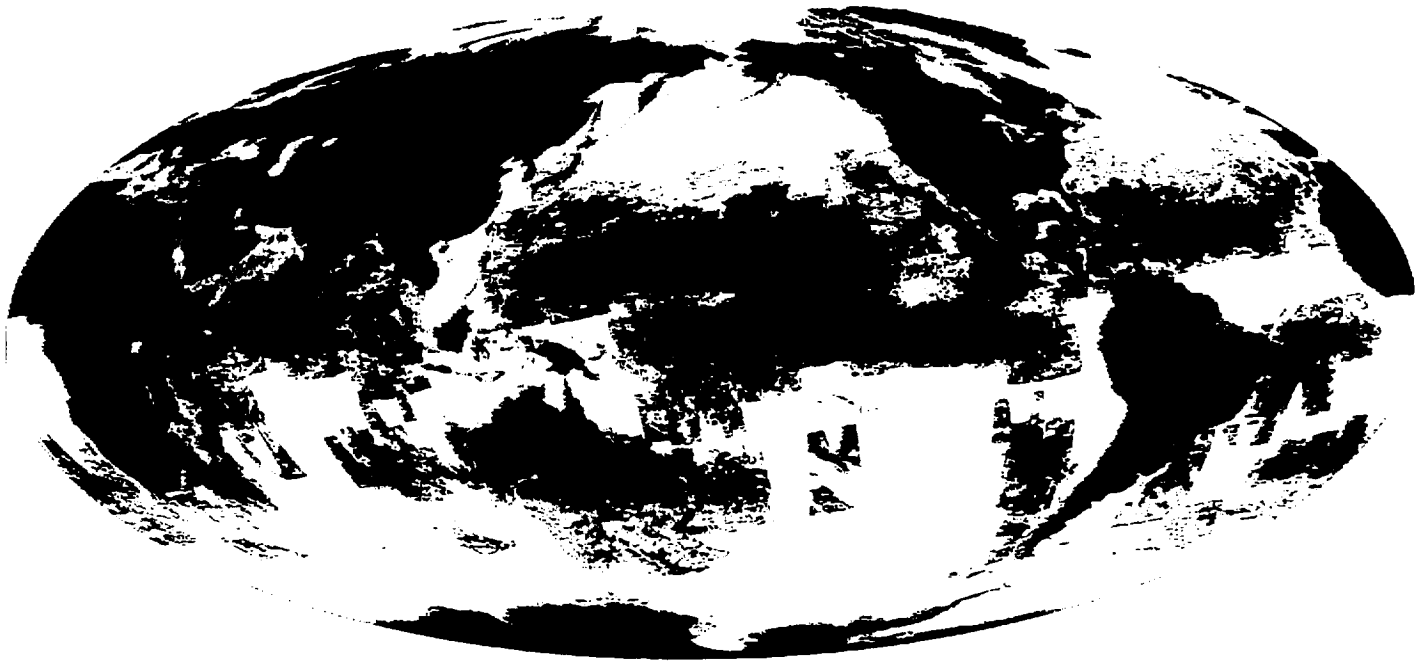
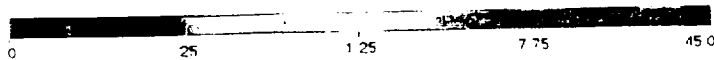
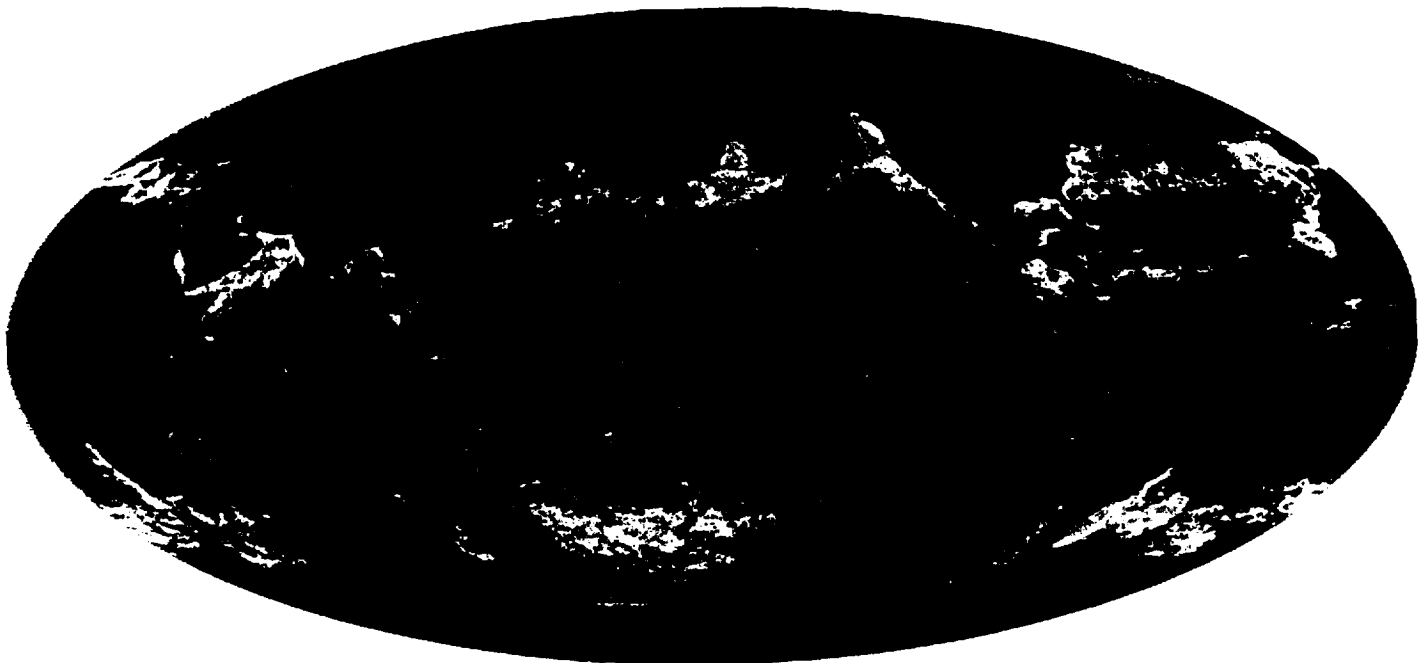


Figure 7. shows input parameters used to to train the neural net. Where C represents Chlorophyll and T for temperature. Each rectangle represents the average pixel for each time slice. Latitude represents the latitude of center images.



CZCS Pigment (mg/m^3) January 1982



Interpolate CZCS Pigment (mg/m^3) January 1982

G. Collins
A. Tren

dClips: A DISTRIBUTED CLIPS IMPLEMENTATION

Y. Philip Li
li@aero.org

Aerospace Corporation

June 30, 1993

525 82
12/1/93
PL

Abstract

A distributed version of the Clips language, dClips, was implemented on top of two existing generic distributed messaging systems to show that: (1) it is easy to create a coarse-grained parallel programming environment out of an existing language if a high level messaging system is used, (2) the computing model of a parallel programming environment can be changed easily if we change the underlying messaging system. dClips processes were first connected with a simple master-slave model. A client-server model with intercommunicating agents was later implemented. The concept of service broker is being investigated.

dClips, was implemented on top of two existing generic distributed messaging systems to show that: (1) it is easy to create a coarse-grained parallel programming environment out of an existing language if a high level messaging system is used as the underlying layer, (2) the computing model of a parallel programming environment can be changed easily if we change the underlying messaging system. In this paper, we describe two versions of dClips implementation on top of two different messaging systems. One messaging system supports only the master-slave model while the other supports a much more flexible communication scheme.

A Master-Slave Model for Task Assignment

dClips was first implemented on top of AERO [Sullivan89], the Asynchronously Executed Remote Operations from UC Berkeley. AERO allows parallel programming in a master-slave mode on a UNIX network. Communication is only allowed between the master and slaves, but not in between slaves. The master process can assign tasks asynchronously, but it has to block and wait for the result to come back.

Introduction

In the process of exploring the opportunities of utilizing multiple workstations on a network as a single parallel computing environment, we have built a simple distributed Clips environment, named dClips, running with multiple parallel Clips processes on a Sun network. Clips, C language integrated production system [Clips91], is a forward-chaining rule-based language with object definition capability. Clips was developed by NASA Johnson Space Center.

As depicted in Figure 1, a single dClips master process controls multiple dClips slave

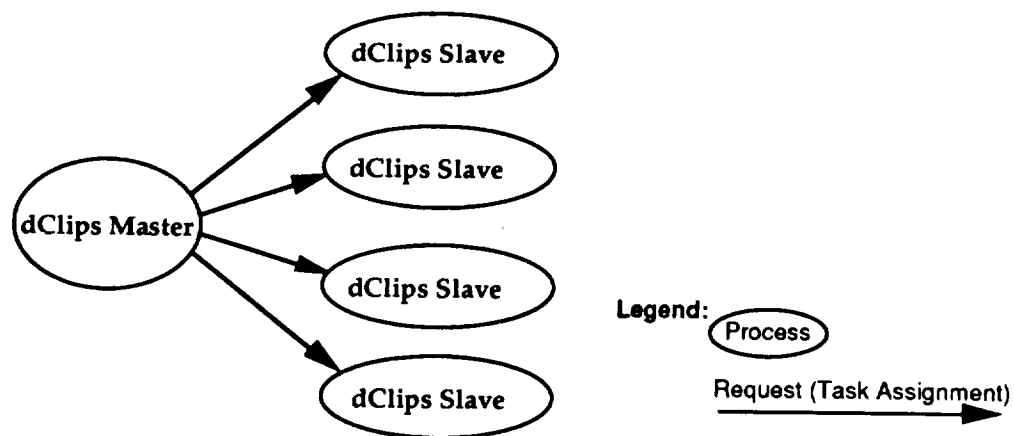


Figure 1. Master-Slave Model for dClips

processes. The master process first asks all the slave processes to load the necessary Clips constructs (i.e., rules, objects, and functions) from the file system into their runtime environments. The master then assigns tasks to slave dClips processes by one of the following three methods:

i). Assert a fact into Clips knowledge base — This request from the master process is executed on all the slave processes simultaneously. If a slave process is busy, it first finishes its current task then asserts the fact. By asserting facts into the working memory of slave processes, the master process could change the inferencing process in the slaves.

ii). Call a Clips function — Any built-in Clips function and user-defined functions in a slave process can be called from the dClips master process. This is a form of remote procedure call in the context of Clips language. Also, the state of the working memory of a dClips slave process can be examined by the master by issuing Clips function call. This allows the master to decide if further task assignment is necessary.

iii). Send a message to a Clips object — A Clips object message can be sent from the dClips master process to dClips slave processes. An active object instance within the slave process can receive messages and process the messages based on the behaviors defined in a message handler.

This version of the dClips implementation is done in C using Clips 5.1. Four function calls are available between the dClips master and slave processes: *loadClipsConstruct*, *assertClipsFact*, *callClipsFunction*, and *sendClipsMessage*. The *loadClipsConstruct* primitive can take a list of construct-files (i.e., a file with Clips rules, objects, and functions) and process them based on the sequence

of the list elements. The sequence in the list is the sequence of execution in the loading process. For example, the list (function.clp object.clp rule.clp) will cause a slave process to load function.clp first, object.clp next, and rule.clp last.

The *assertClipsFact* primitive takes a string with a single Clips fact and requests every slave to assert it into the knowledge base. The *callClipsFunction* primitive takes a string with a single function name and the function parameters, and sends it to all the slave processes.

The *sendClipsMessage* primitive is also capable of passing a list of messages from master to slave. Each message is itself a list in the form: (class-name instance-name method-name method args). The slave that receives a *sendClipsMessage* request processes the messages based on the sequence of the list elements. This allows multiple class methods to be defined and executed in sequence as a single work assignment.

An Application

A Clips-based image data access application, DataHub [Handley92], has been ported to the dClips/Aero environment. The master process issues concurrent image data access/conversion requests for different dataset types. Each dataset type has different data format and data semantics. The knowledge about the image datasets is stored in Clips constructs, and loaded by the slave processes at startup time. One dataset type is handled by one slave process. There is no interaction needed among slave processes. Data conversion tasks are both CPU intensive due to format changes (e.g., byte swap, data decompression) and I/O extensive due to massive read and write of files.

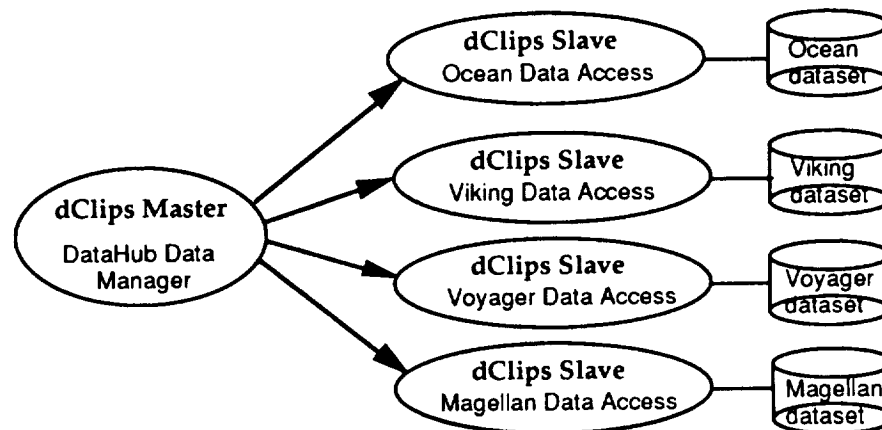


Figure 2. DataHub Data Manager on top of dClips

The DataHub data manager with the same master-slave task assignment scheme has also been ported to the dClips/ISIS environment (see next section for details). The master slave model stays with the ISIS implementation because of the nature of the application, rather than the limitation of the ISIS computing model.

A Client-Server Model with Service Brokers

The master-slave process model imposed by AERO is not desirable if we want to build systems with communicating intelligent agents. After evaluating Sun's ToolTalk™ [ToolTalk91] and Cornell's ISIS [ISIS90] for an alternative distributed computing model, we decided to build dClips on top of ISIS. ToolTalk was not chosen because: 1) the message arrival sequence from multiple senders in a network environment is not guaranteed, 2) only 1 handler is allowed for a request message (others are observers). Message arrival sequence is important because the arrival sequence of messages for asserting a fact or for updating object instances in the dClips environment is critical to the local Clips inferencing process. Different message arrival patterns could result in different inference outcomes. Furthermore, the constraint of having a single handler for a request message makes it unnatural for task distribution/assignment in a parallel programming environment.

On the other hand, ISIS, developed at Cornell University, provides a set of tools built around virtually synchronous process groups and reliable group multicast [Birman91]. A virtually synchronous distributed system has the following characteristics: (1) all processes observe events in the same order (global order and causality), (2) an event notification is delivered to all or none of the audience (atomicity). A virtually synchronous system looks synchronous to every process in the system, but executes asynchronously. For the dClips implementation, the virtually synchronous broadcast (cbcast, for causal broadcast), which guarantees the causality and atomicity, was the main reason for using ISIS as the underlying distributed computing model.

Figure 3 shows the architecture of ISIS-based dClips, where a set of dClips Server processes team up with a dClips Administrator process to form a process group. This process group provides the cooperative problem solving capability to the outside world. The dClips Administrator plays the role of a *service broker*, providing a consistent interface to the

outside clients, while the details of the server processes are transparent to the clients. The interface between a service broker and its clients has yet to be defined. At this point, the CORBA (Common Object Request Broker Architecture) IDL (interface definition language) type interface is being considered [OMG91]. The interface between the dClips Administrator and the dClips Servers is a shared knowledge base with a set of common access methods.

The dClips Servers form a conceptual hierarchy, which is known to the dClips world, but is not visible to the ISIS environment. In other words, this Server hierarchy is not a hierarchy of ISIS process groups. In the ISIS environment, all the servers are equal members of a single process group. Broadcasts to the group will reach every server process in the same order. Each server is an autonomous problem solving agent with its own knowledge base and its own task. The Server hierarchy defined within dClips environment helps a server to find another potential problem solver if a problem cannot be solved locally.

A shared knowledge base is available to dClips Servers for knowledge exchange and interaction, which is designed to facilitate the cooperative problem solving process conducted by multiple dClips Servers. At the same time, each dClips server can have its own individual non-shared knowledge base. The Server hierarchy is defined as a Clips Class Hierarchy within the shared knowledge base, which is known to every server. The message communication between servers can be: (1) a broadcast to the whole process group, or (2) a message to a designated server.

The shared knowledge base is realized by having a set of Clips constructs replicated in each server. Each server loads in this shared knowledge base at initialization time. Any update to any of the objects in this shared knowledge base in any dClips Server will trigger a broadcast of the update to other members in the process group. A server applies the updates sent in from other servers one by one as if they are local updates to the knowledge base. Since the shared knowledge is designed to keep only the critical knowledge that needs to be shared among servers, the size of this shared knowledge base should be small. The effort to keep it consistent across multiple servers, i.e., sending and receiving update messages and applying updates triggered by remote update messages, should be minimal.

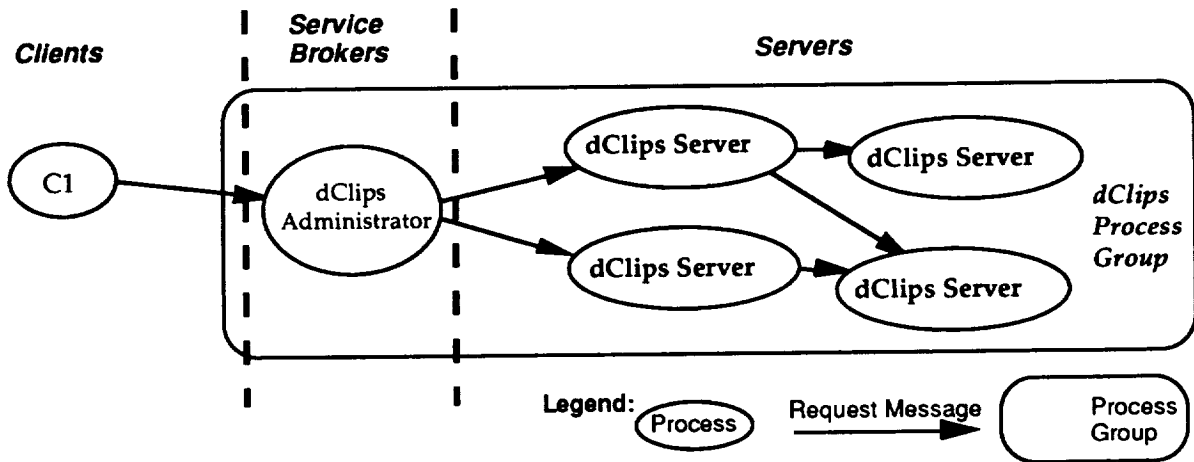


Figure 3. Architecture of ISIS-based dClips

Future Opportunities

Based on the client-server model of dClips, we would like to pursue the following extensions:

i). dClips with Database Access Capability — This involves a dClips database gateway, which runs as a database client to some database server. An intelligent agent can not be intelligent without necessary knowledge about the real world. Accessing existing databases is one way of acquiring data/knowledge from the outside world. As shown in Figure 4, a database pass-through process can serve as

the gateway to the database server. The function of this gateway can be as simple as passing a SQL statement to a relational database system and receiving the results back in a buffer. Or it can provide more sophisticated functions such as allowing joins of tables across multiple database systems.

ii). A distributed blackboard system on top of dClips — The ISIS-based dClips implementation can easily evolve into a distributed blackboard system. This can be done by making dClips server processes run as *knowledge sources* in a blackboard system and by

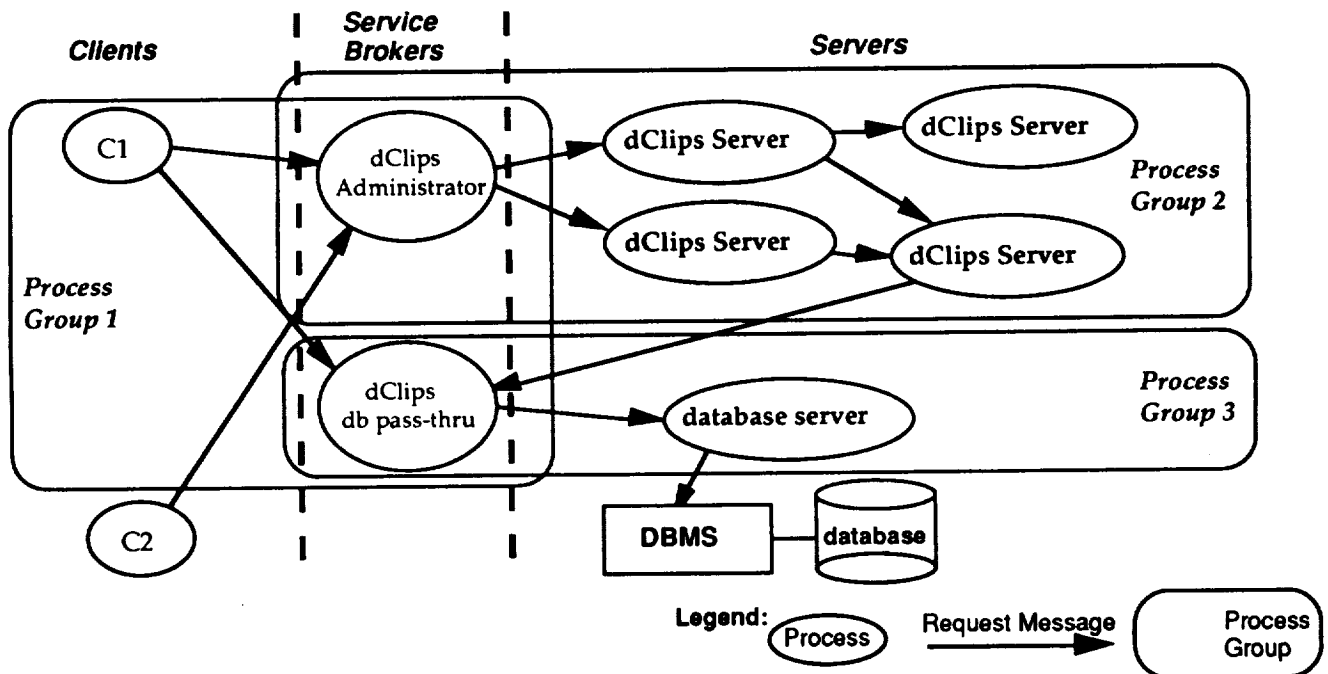


Figure 4. dClips with Database Access Capability

using the shared knowledge base among dClips servers as the *blackboard* [Nii86]. A blackboard system like this is a realization of the original blackboard metaphor because there is no centralized control mechanism involved in the blackboard reasoning process. Each knowledge source reacts only to the change on the blackboard. A domain problem can be solved cooperatively this way by multiple knowledge sources.

References

- [Birman91] Birman, K., Schiper, A., Stephenson, P., Lightweight Causal and Atomic Group Multicast, ACM Transactions on Computer Systems, Vol. 9, #3, August 1991.
- [Clips91] Clips, Version 5.1, Reference Manual, Software Technology Branch, Johnson Space Center, Sep. 1991.
- [Handley92] Handley, T., Li, Y. P., DataHub: Knowledge-based Data Management for Data Discovery, ISY Conference on Earth and Space Science Information Systems, Feb. 10-13, 1992.
- [ISIS90] ISIS, Version 2.1, User's Guide and Reference Manual, Cornell University, Sep. 1990.
- [Nii86] Nii, H. P., Blackboard System: The Blackboard Model of Problem Solving and the Evolution of Blackboard Architectures, THE AI Magazine, Summer, 1986, p. 38-53.
- [OMG91] Object Management Group and X/Open, Common Object Request Broker: Architecture and Specification, December 1991.
- [Sullivan89] Sullivan, M., Anderson, D., Marionette (Also Known As Aero): a System for Parallel Distributed Programming using a Master Slave Model, IEEE 9th International Conference on Distributed Computing Systems, 1989.
- [ToolTalk91] TookTalk, Version 1.0, Programmer's Guide, SunSoft, Dec. 1991.

DATA VISUALIZATION AND SENSOR FUSION
 Characteristics of Remote Sensing Data- DMSP

HUGHES

UNIVERSITY OF CHICAGO
 FRANKLIN MEDICAL INSTITUTE

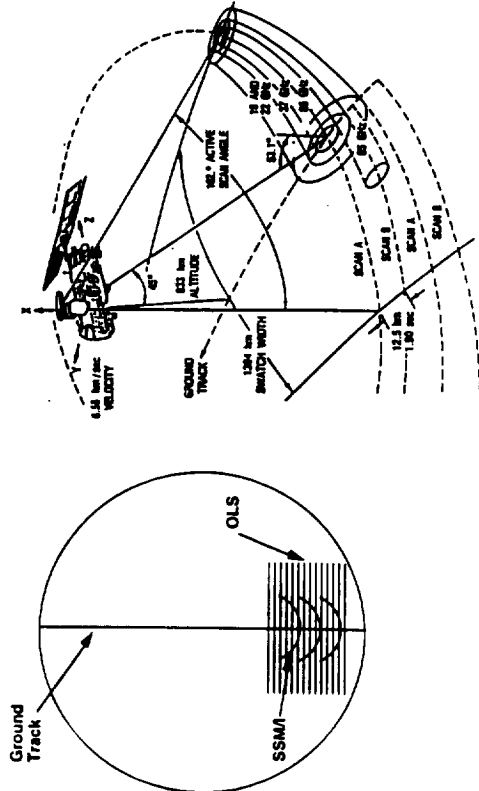
SSM/I		OLS	
Freq (GHz)	Resolution (km)	Visible	IR
19.1 V/H	50	0.4 - 1.1 μ m	0.5 km/ 2.5 km
22.2 V	25	10 - 12 μ m	0.5 km/ 2.5 km
37 V/H	25		
85.5 V/H	12.5		

Applications:	Applications:
Ocean Surface Wind Speed (19, 22, 37)	Visible: Clouds, Features, etc
Rain Rate (85)	IR: Temperature, Cloud Height
I.c.e (37)	
Water Vapor (22)	

DATA VISUALIZATION AND SENSOR FUSION
 Characteristics of Remote Sensing Data- DMSP

HUGHES

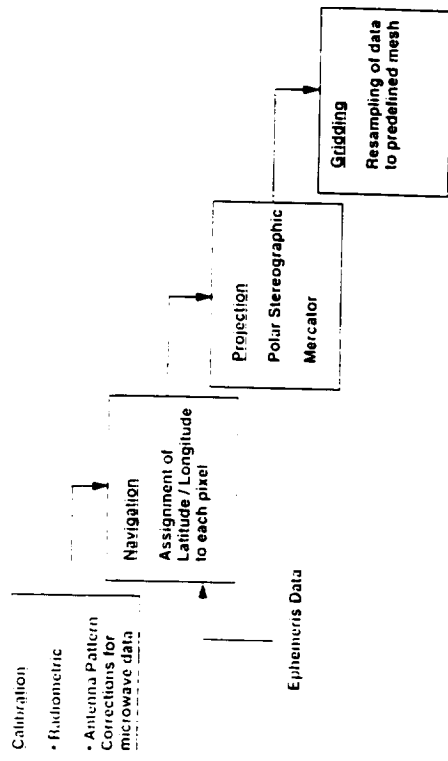
UNIVERSITY OF CHICAGO
 FRANKLIN MEDICAL INSTITUTE



DATA VISUALIZATION AND SENSOR FUSION
 Calibration, Navigation, Projection and Gridding

HUGHES

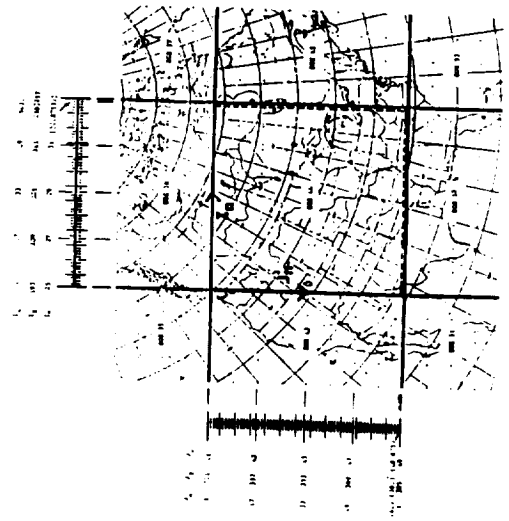
UNIVERSITY OF CHICAGO
 FRANKLIN MEDICAL INSTITUTE



DATA VISUALIZATION AND SENSOR FUSION
 Gridding of DMSP Data to Standard Meshes

HUGHES

UNIVERSITY OF CHICAGO
 FRANKLIN MEDICAL INSTITUTE



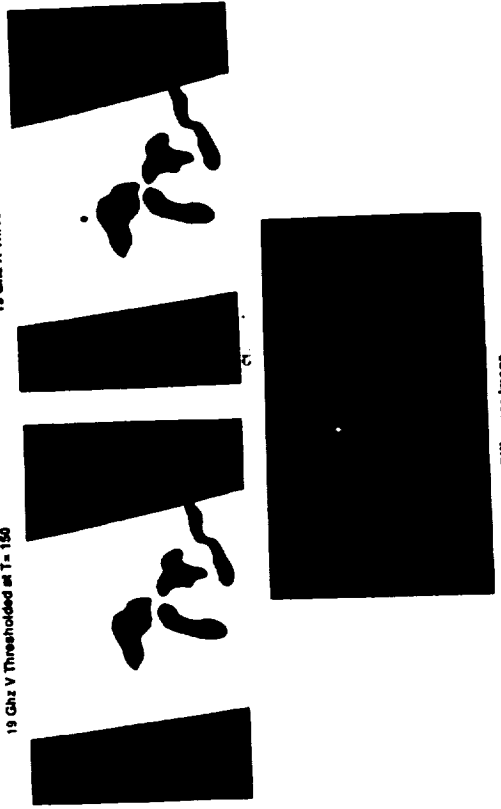
- Medical imaging software not as well suited to projection and gridding problems as anticipated
- Raw data preprocessed prior to delivery to Univ of Chicago
- All data referenced to same grid for initial convenience
- Current data is in 512 by 512 pixel files

- Operating System- Unix
- Programming Language- C
- Application Package- IDL
- Data Format(s)- Binary (Can convert to other formats)

- Image Matching Based on Landmarks
- Contouring and Thresholding
 - Threshold Settings
 - Least Squares Fits Between Images
 - Noise and Contrast
 - Image Characteristics
 - Edge Sharpness of ROI's
 - Resolution

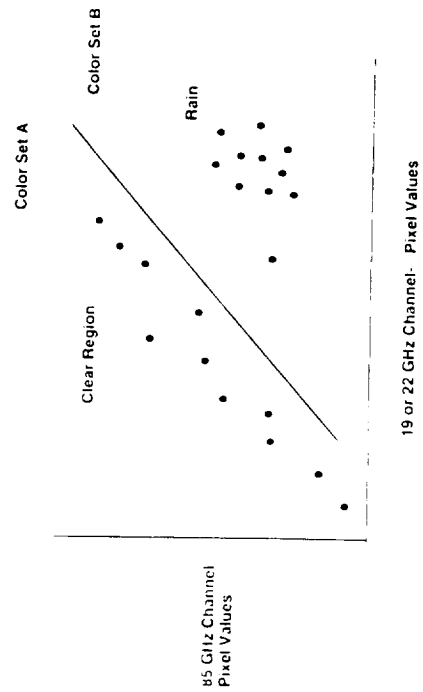
19 Ghz V Thresholded at T= 150

19 Ghz H Thresholded at T= 150

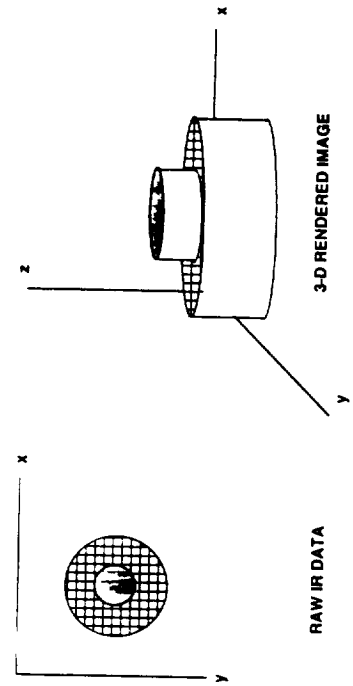


- User Interface and Manual Control Important
- Operator Interaction Capability Necessary to:
 - Select Regions of Interest
 - Adjust Parameters (e.g. Thresholds)
 - Perform Initializations (For least- squares techniques)
- Color:
 - Not as important as initially presumed in medical images.
- Need to evaluate for remote sensing images

- Motivation- Correlation between images can be used to define color table which will highlight phenomena of interest
- At a given pixel location the values of the pixels from two images are used as "lookup" pointers to a color table
- Issues- Defining features and selecting colors



Altitude of Pixels in Thermal IR Images Can be Estimated From Their Temperature
 Objective- Apply 3-D Medical Imaging Techniques to IR Images



- Work done at Hughes shows:

Results of principal components transformation of microwave channels can extract useful data

Principal components can be used to drive hue, saturation, and intensity to produce useful visualizations of multi channel data

Koscielny, A. J. and Savage, R. C. "A Color Composite Display for SSM/I Imagery", 5th Annual Conference on Information and Processing Systems for Meteorology, Oceanography, and Hydrology, Anaheim 1/89

- Apply Visualization Algorithms to Remote Sensing Data
- Acquire Additional Data for Evaluation
- Evaluate Visualization Results as a Function of :
 - Image Characteristics and Desired Information
 - Image Processing and Visualization Parameters
- Investigate Applications of Principal Component Techniques
- Three- Dimensional Visualization of IR Images



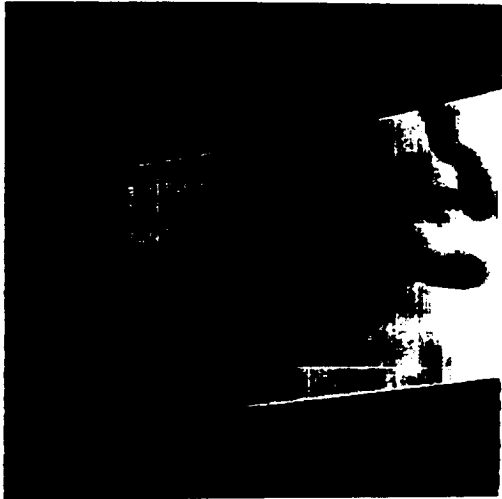
DATA VISUALIZATION AND SENSOR FUSION
Current Data Set- SSM/ 19 GHz H

HUGHES
UNIVERSITY OF CHICAGO
FRANKLIN MEDICAL INSTITUTE



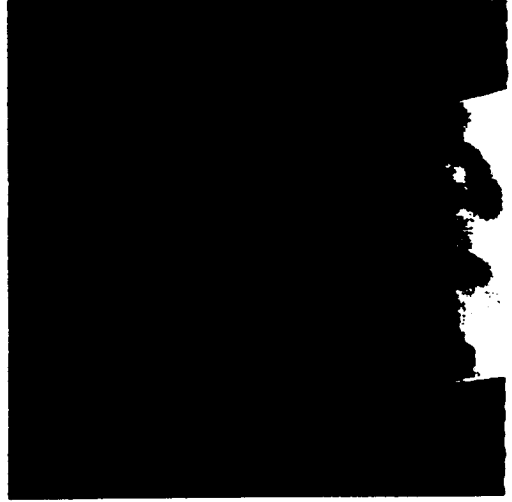
DATA VISUALIZATION AND SENSOR FUSION
Current Data Set- SSM/ 19 GHz V

HUGHES
UNIVERSITY OF CHICAGO
FRANKLIN MEDICAL INSTITUTE



DATA VISUALIZATION AND SENSOR FUSION
Current Data Set- SSM/ 37 GHz V

HUGHES
UNIVERSITY OF CHICAGO
FRANKLIN MEDICAL INSTITUTE



DATA VISUALIZATION AND SENSOR FUSION
Current Data Set- SSM/ 22 GHz V

HUGHES
UNIVERSITY OF CHICAGO
FRANKLIN MEDICAL INSTITUTE



HUGHES

UNIVERSITY OF CHICAGO
FRANKLIN MEDICAL INSTITUTE

DATA VISUALIZATION AND SENSOR FUSION

Current Data Set- SSM/I 37 GHz H



HUGHES

UNIVERSITY OF CHICAGO
FRANKLIN MEDICAL INSTITUTE

DATA VISUALIZATION AND SENSOR FUSION

Current Data Set- SSM/I 85 GHz V



HUGHES

UNIVERSITY OF CHICAGO
FRANKLIN MEDICAL INSTITUTE

DATA VISUALIZATION AND SENSOR FUSION

Current Data Set- SSM/I 85 GHz H



DATA COMPRESSION

**Dr. James A. Storer
Brandeis University**

August 5, 1993

Data Compression

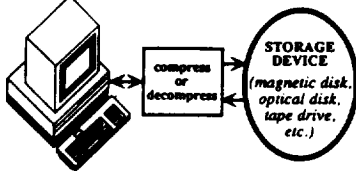
James A. Storer
Computer Science Dept.
Brandeis University

Joint Work With:

- B. Carpentieri* (Ph.D. Student)
- M. Cohn* (Faculty)
- C. Contantinescu* (Ph.D. Student)
- S. De Agostino* (Ph.D. Student)
- E. Lin* (Post-Doc)
- Q. Ye* (Ph.D. Student)
- R. Zito-Wolf* (Ph.D. Student)

Applications of Data Compression

1. Data Storage



2. Data Communication



3. Machine Learning

lossless = decompressed data
is identical to the original

lossy = decompressed data
may be an approximation to the original

Key types of data:

- text
- computer source/object code
- data bases
- numerical data
- speech
- music
- gray-scale images
- color images
- graphics
- CAD data
- animation
- half-tone/fax data
- finger print images
- bank check images
- map and terrain data
- medical imagery
- scientific and instrument data, space data
- image sequences
- video

Examples of Speeds Required for Real Time Processing:

Text sent over a modem ~ 2,400 bits per second
 (Depending on the cost of the modem, commonly used speeds range from 1,200 bits per second to 9,600 bits per second.)

Speech ~ 100,000 bits per second
 (One government standard uses 8,000 samples per second, 12 bits per sample.)

Stereo Music ~ 1.5 million bits per second
 (A standard compact disc uses 44,100 samples per second, 16 bits per sample, 2 channels.)

Picture Phone ~ 12 million bits per second
 (A low resolution black and white product might require 8 bits per pixel, 256x256 pixels per frame, 24 frames per second.)

Black&White Video ~ 60 million bits per second
 (A medium resolution product might use 8 bits per pixel, 512 by 512 pixels per frame, 30 frames per second.)

HDTV ~ 1 billion bits per second
 (A proposed standard has 24 bits per pixel, 1024 by 768 pixels per frame, 60 frames per second.)

Data Compression Research at Brandeis

Lossless Compression:

- Systolic Algorithms
- High Speed Hardware
- Optimal Poly-Log Algorithms for Parallel Machines
- Poly-Log Approximation Algorithms for Simple Architectures
- Complexity of Off-Line Vs On-Line Encoding

Image Compression:

- Tree-Structured Vector Quantization
- On-Line Adaptive VQ with Variable-Sized Vectors
- Visualization Tools
- Fast "Browsing" of Scientific Images
- Applications to Scene Analysis and Object Classification

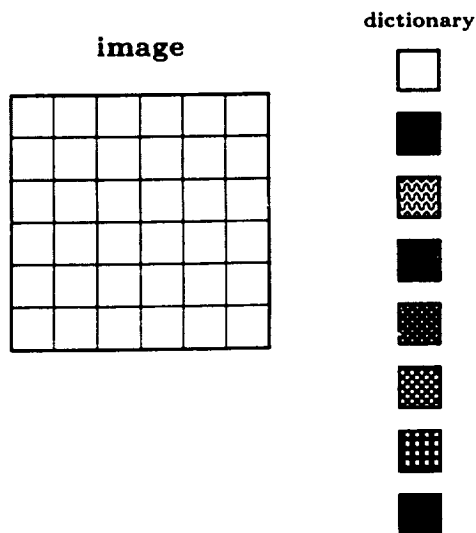
Video Compression:

- Real-Time Displacement Estimation with Variable-Size Blocks
- Sub-linear Algorithms
- Hardware Design for Integrated Systems
- Applications to Machine Learning

Error Resilient Compression

Image Compression with Vector Quantization

IDEA: Map sub-arrays of pixels ("vectors") to the "closest" vector in a dictionary.



On-Line Adaptive VQ

Main idea:

Process an image in a single pass by adaptively constructing a codebook of variable size rectangles.

No training is used; "learning" is done on-line in a fashion analogous to LZ methods for lossless text compression.

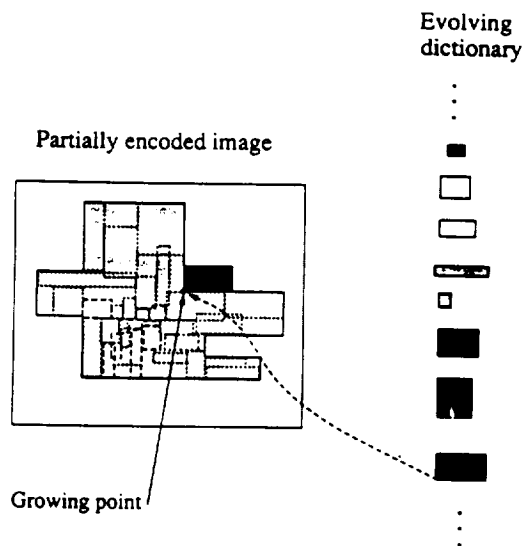
Example of the covering pattern



Decompressed image

Covering pattern

On-Line Adaptive VQ



Generic Encoding Algorithm

1. Initializations.

- Dictionary $D = 256$ single byte values
- Growing Points Pool GPP = one pixel

2. Repeat until there are no more growing points in GPP:

- (a) *{ Growing Heuristic }*
Choose a growing point GP from GPP.
- (b) *{ Match Heuristic }*
Find the "best" match block b .
Transmit $\lceil \log_2 |D| \rceil$ bits for the index of b .
- (c) *{ Update Heuristics }*
Update D and GPP.
(Delete if necessary).

Generic Decoding Algorithm

1. *{ perform Step 1 of the encoding algorithm }* (Initializations)

2. Repeat until there are no more growing points in GPP:

- (a) *{ Perform Step 2a of the encoding algorithm }*
(Choose a GP)
- (b) Receive $\lceil \log_2 |D| \rceil$ bits - the index of block b .
Retrieve b from D and output b at the position determined by GP.
- (c) *{ Perform Step 2c of the encoding algorithm }*

Growing Strategies

Wave:



Diagonal:



Lifo:

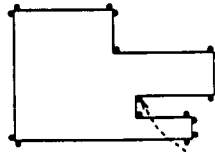


* Circular: (Used for all experiments)



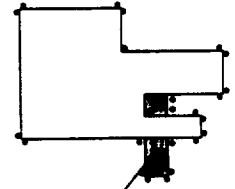
Growing Points

Partially encoded image



Selected growing point

Partially encoded image
after adding a new block



New matched block

The Match Heuristic

Decides what block (entry) from the dictionary best matches the area of the image originating in the selected GP

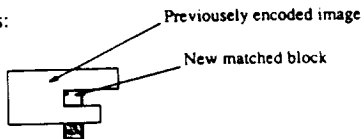
Parameters:

- Distortion measure (eg. L2)

- The threshold

- Overlapping blocks:

- FIRST cover



- LAST cover



* - AVERAGE (Used for all experiments)



MATCH HEURISTIC cont. :

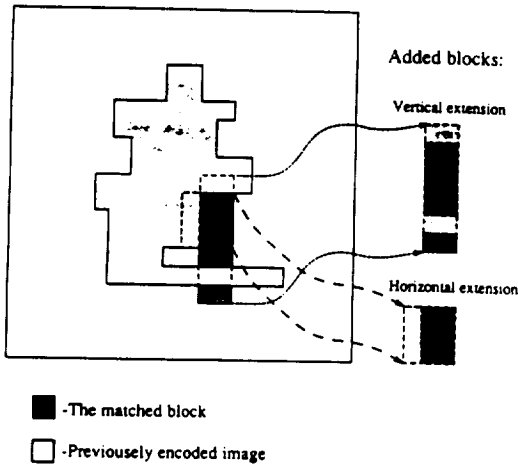
GREEDY: Chooses the biggest block satisfying the threshold.

* **Moderated_GREEDY:** (Used for all experiments)

Best match must be significantly larger than rivals of almost equal quality.

Dictionary Update Heuristic

* One-Row plus One-Column (Used for all experiments)



Deletion Heuristic

FREEZE: Once the dictionary is full, "freeze" it (i.e. do not allow any further entries to be added).

LRU: Delete the entry that has been least recently used.

* **FIFO:** (Used for all experiments)
Keep the dictionary entries in a queue implemented as a circular array.

Test Images

BrainCat: Cat-scan brain image, 512 by 512 pixels, 8 bits per pixel.

BrainMrSide: Magnetic resonance medical image that shows a side cross-section of a head, 256 by 256 pixels, 8 bits per pixel; this is the medical image used by Gray, Cozman, and Riskin [1991,1992].

BrainMrTop: Magnetic resonance medical image that shows a top cross-section of a head, 256 by 256 pixels, 8 bits per pixel.

NASA5: Band 5 of a 7-band image of Donaldsonville, LA; the least compressible of the 7 bands by the UNDX compress command.

NASA6: Band 6 of a 7-band image of Donaldsonville, LA; the most compressible of the 7 bands by the UNDX compress command.

WomanHat: The standard woman in the hat photo, 512 by 512 pixels, 8 bits per pixel.

LivingRoom: Two people in the living room of an old house with light coming in the window, 512 by 512 pixels, 8 bits per pixel.

FingerPrint: An FBI finger print image, 768 by 768 pixels, 8 bits per pixel; includes some text at the top.

HandWriting: The first two paragraphs and part of the figure of page 165 of *Image and Text Compression* (Kluwer Academic Press, Norwell, MA) written by hand on a 10 inch high by 7.5 inch wide piece of gray stationary scanned at 128 pixels per inch, 8 bits per pixel; approximately 1.2 million bytes.

COMPARISON OF COMPRESSION RATIOS FOR THE SAME SNR (PSNR)

	"VERY GOOD"		"GOOD"		"FAIR"	
	SNR (PSNR)	COMPRESSION OURS/SPEC	SNR (PSNR)	COMPRESSION OURS/SPEC	SNR (PSNR)	COMPRESSION OURS/SPEC
BrainCat	29(36)	4.2/2.9	22(29)	8.7/4.8	17(25)	12.5/6.7
BrainMrSide	29(37)	4.0/4.6	27(37)	4.8/5.4	21(31)	10.8/15.8
BrainMrTop	27(35)	2.8/2.5	21(29)	5.6/3.6	15(23)	10.8/6.6
NASA5	31(41)	4.1/4.1	28(39)	5.5/5.9	26(37)	7.4/8.5
NASA6	46(52)	22.2/2.4	41(47)	79.4/66.7	37(45)	97.7/65.1
WomanHat	32(36)	6.5/8.8	30(35)	9.2/13.3	27(33)	14.1/22.4
LivingRoom	30(36)	5.1/5.9	27(33)	7.4/9.0	25(31)	10.5/16.0
FingerPrint	32(35)	6.2/6.3	24(27)	22.8/26.5	22(25)	32.3/34.3
HandWriting	32(33)	16.5/9.6	25(26)	59.2/36.7	17(18)	175.0/62.3

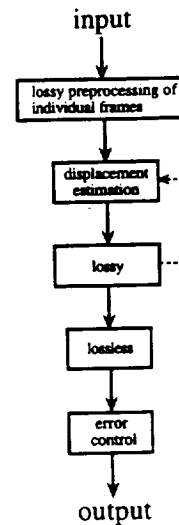
ALL ENTRIES USE: { WAVE
L2 - 4x4
AVERAGE
GREEDY

COMPARISON OF COMPRESSION RATIOS FOR FULL SEARCH VS KD-TREE

NAMES	"Very Good"		"Good"		"Fair"	
	SNR	Compression	SNR	Compression	SNR	Compression
	FULL / TREE	FULL / TREE	FULL / TREE	FULL / TREE	FULL / TREE	FULL / TREE
BrainCAT	29 (29)	4.2 / 4.2	22 (22)	8.9 / 8.8	18 (18)	12.5 / 12.6
BrainMR_Side	29 (28)	4.0 / 4.1	27 (26)	4.8 / 4.9	21 (21)	10.4 / 10.4
BrainMR_Top	27 (27)	2.8 / 2.8	21 (21)	5.6 / 5.5	15 (15)	10.8 / 10.5
NASA5	31 (30)	4.1 / 4.3	28 (28)	5.5 / 5.5	26 (26)	7.4 / 7.5
NASA6	46 (46)	22.2 / 22.7	41 (41)	79.4 / 82.1	39 (40)	97.7 / 105.6
WomanHat	32 (32)	6.5 / 6.4	30 (30)	9.2 / 8.9	27 (27)	14.1 / 14.1
LivingRoom	30 (29)	5.1 / 5.2	27 (27)	7.4 / 7.3	25 (25)	10.5 / 10.5
FingerPrint	32 (32)	6.2 / 6.2	24 (24)	27.8 / 27.6	22 (22)	38.3 / 37.8
HandWriting	32 (32)	16.5 / 16.5	25 (25)	59.2 / 60.8	17 (18)	175.0 / 170.7

ALL ENTRIES USE: $\left\{ \begin{array}{l} \text{WAVE} \\ \text{L2-4x4} \\ \text{AVERAGE} \\ \text{GREEDY} \end{array} \right.$

Simplified Video Compression System



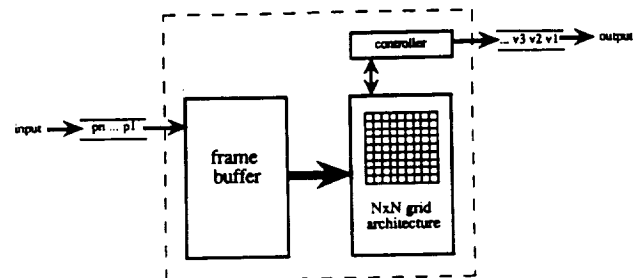
Displacement Estimation

Idea : Approximate interframe motion by piecewise translation of blocks of pixels.

(Rotation, zooming, etc., approximated by block translation, if blocks are small.)

Note : Displacement estimation is a crucial part of the MPEG standard .

The Model of Computation



DISPLACEMENT ESTIMATION ENCODER

- Input/Output is serial.
- n : number of pixels per frame.
- Each processor corresponds to a block of k pixels (i.e. $n/k = N \times N$).
- Controller communicates with only one processor.
- Data for the current and previous frame is processed by the grid while data for the next frame is filling up the frame buffer

Superblocks

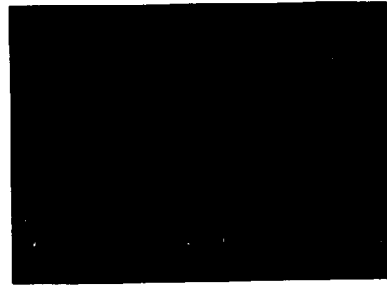
Def. : Superblock at time t : set of adjacent blocks with the same DMD at time $t-1$

Properties of Superblocks:

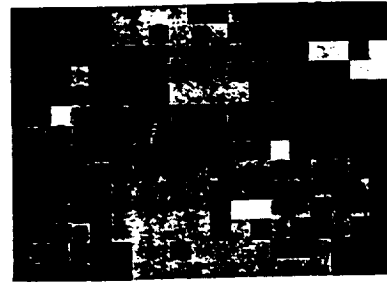
- Superblocks will represent areas of the image with the same displacement vector
- Superblocks may have no prescribed shape
- Superblocks may grow and shrink from frame to frame

Idea : Use a parallel grid architecture to segment each frame into superblocks

Note : We will not need the monotonicity assumption

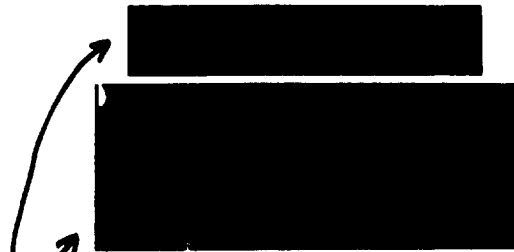


a) KIDS2 b) KIDS3 c) Prediction of KIDS3, block size 4 d) prediction of KIDS3, block size 8



Segmentation of the frame KIDS3 into superblocks, the initial block size is 16

SCIENTIFIC IMAGE SEQUENCES



EXAMPLES (PROVIDED BY M. BOTTS)

DAILY TEMPERATURE AND DATA

3D RADAR

* OFTEN THESE SEQUENCES "LOOK" LIKE VIDEO AND DISPLACEMENT ESTIMATION TECHNIQUES ARE EFFECTIVE

PRELIMINARY RESULTS:

SUPERBLOCKS GIVE A FACTOR OF 3 TO 4 IMPROVEMENT OVER FIXED SIZE BLOCKS.

CONJECTURE:

ANOTHER FACTOR OF 2 CAN BE GAINED FROM ADAPTIVELY TREATING SPLIT REGIONS.

GOAL:

REAL-TIME HIGH FIDELITY VIDEO AND IMAGE SEQUENCE COMPRESSION OF OVER 1000-TO-1.

**SAMS: A SPATIAL ANALYSIS & MODELING
SYSTEM FOR ENVIRONMENTAL MONITORING**

**Mr. Fran Stetina
Goddard Space Flight Center**

August 5, 1993

SAMS

A Spatial Analysis & Modeling System
for
Environmental Monitoring

Co - Pi Fran Stetina
Co - Pi John Hill

OBJECTIVE: Develop an end - end system to
support environmental monitoring
and management

→ **Towards an EOS pilot Direct Readout
Ground processing & analysis system**

- Real Time data reception
- Multi-discipline Analysis & Modeling
- Product development (expert systems)
- Information Distribution

Approach: Integrate software modules into
a single computer environment

Data Acquisition

Variable data rates
Real time acquisition
Multi-satellite/sensors

Multi Discipline

ATMOSPHERIC PROCESSES
LAND PROCESSES
OCEAN PROCESSES

Approach : Integrate software modules into a single computer environment

PRODUCT DEVELOPMENT & DISPLAY

- MARINE RESOURCES
- WATER RESOURCE MANAGEMENT
- DISASTER MANAGEMENT
- AGRICULTURE
- ENVIRONMENTAL MONITORING & MANAGEMENT

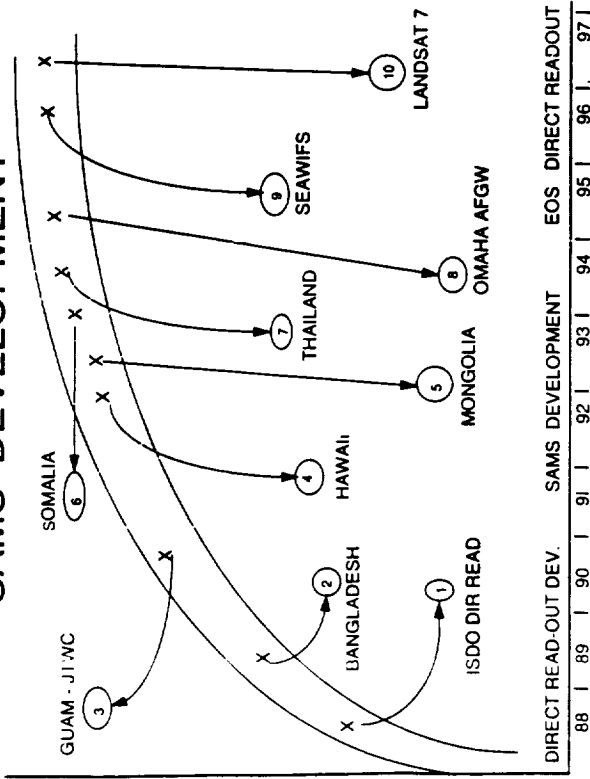
DATA ARCHIVING & DISTRIBUTION

- DATA BASE MANAGEMENT
- DATA FUSION
- INTERNET
- PEACESAT
- EARTH ALERT
- NASA ACTS
- HARD COPY & STORAGE

SAMS SYSTEMS INSTALLED

- 1. ISDO DIRECT READ-OUT** TIROS N POLAR ORBITER
IVAS IMAGE PROC.
2 UVAX - 2 's & 1 UVAX 4000
- 2. BANGLADESH** GMS, POLAR ORBITER & ARGOS - DCP's
2 UVAX 3400, IVAS IMAGE PROC. & P.C.'s
- 3. GUAM - JTWC** GMS, TIROS, DMSP, SSMT, SSMT & GTS
3 UVAX 3400, 2 - IVAS IMAGE PROC. & P.C.
MAC IIFX
- 4. HAWAII** GMS DISTRIBUTION & LOCAL DISPLAY
1 UVAX 4000 & MAC for DISPLAY
- 5. MONGOLIA** TIROS GEOSPHERE BIOSPHERE PROJECT
2 UVAX 2'S, IVAS & MAC for DISPLAY
- 6. SOMALIA** METOSAT DISTRIBUTION & LOCAL DISPLAY
1 UVAX 2 & MAC IIFX
GMS & NEXRAD RADAR DATA
UVAX 3400 & QUADRA 800
- 7. THAILAND**

SAMS DEVELOPMENT



Systems under development

- H.A.R.C.** REMOTE SENSING / G I S LAB
- 8. OMAHA AFGW** METEOSAT & GMS DATA FROM NASA
GOES NEXT SYSTEM
UVAX & MAC IIFX FOR DISPLAY
D.E.C. ALPHA CPU
- 9. SEAWIFS** INDEPENDENT FRAME FORMATTER
SILICON GRAPHICS CPU
- 10. LANDSAT 7** D.E.C. ALPHA CPU INCLUDING ALL
IMAGE PROCESSING FUNCTION

LAND-SURFACE TESTBED FOR EOSDIS

**Dr. William Emery
University of Colorado**

August 5, 1993

LAND-SURFACE TESTBED FOR EOSDIS

Bill Emery University of Colorado
Jeff Dozier University of California - Santa Barbara
Tim Kelley University of Colorado
Dan Baldwin University of Colorado
Paul Rotar NCAR Mesa Facility

August 5, 1993

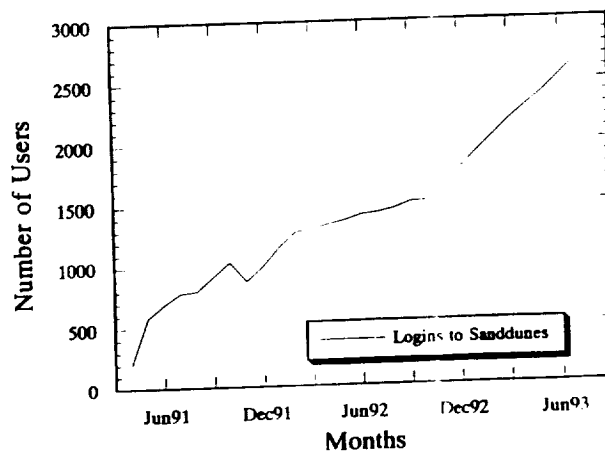
Historical Background on Sanddunes

- To provide a pre-processed 640 x 555 non-calibrated image centered over Colorado via ftp to end users free of charge.
- Image size increased to include great plains region 1240 x 644.
- Users needed On-Line-Browse to preview the images for cloud and area coverage.
- Users need viewing software, NDVI, Cloud filtering software
- West coast was added to the image increasing the size to 1640 x 1544 the images were now 3.9 megabytes per channel
- Allow users to navigate (geo-register) images to their coordinates pick the resolution, projection along with several other options.
- Added GOES images to the data system

Work In Progress

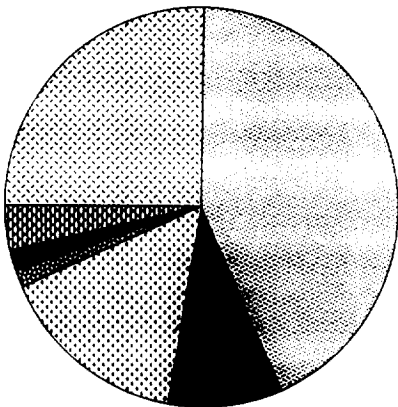
- Development of a stand-alone data archive system
- Addition of DOMSAT data
- Addition of GRADS overlays

Total Number of Registered Users



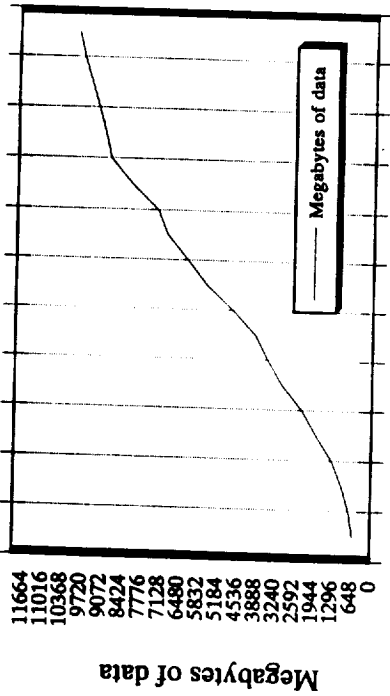
Type Of Users By Category

- Government
- College
- Highschool
- JrHighschool
- Company
- Foreign Gov
- Foreign School
- NonListed

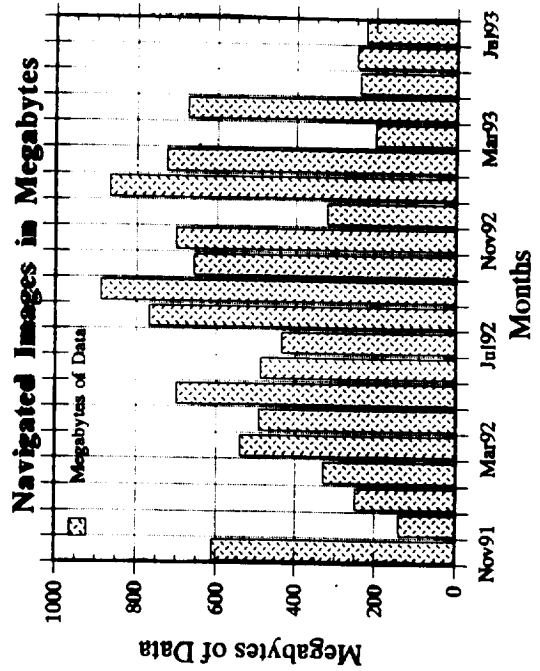


Individual Users

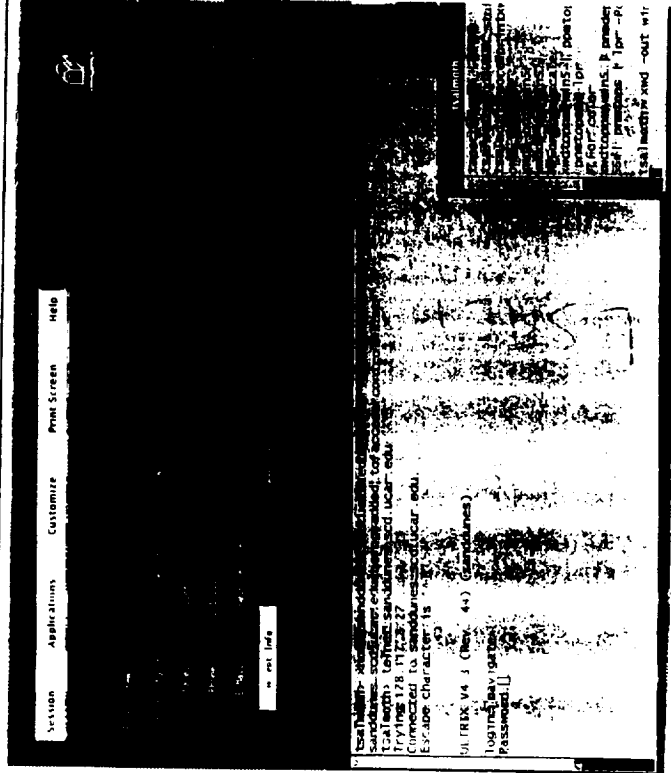
Total Navigated Images In Megabytes

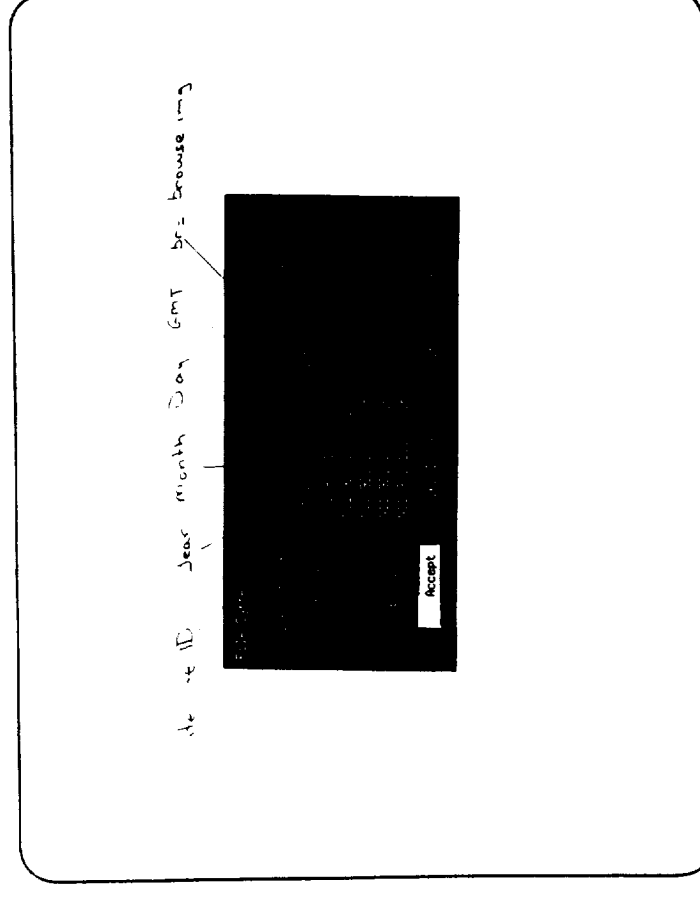
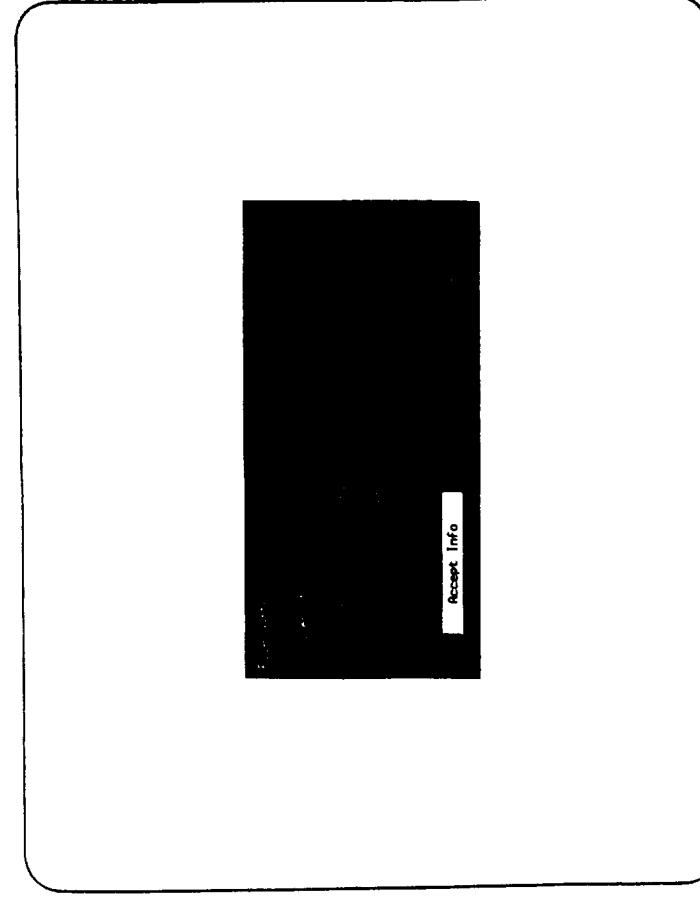
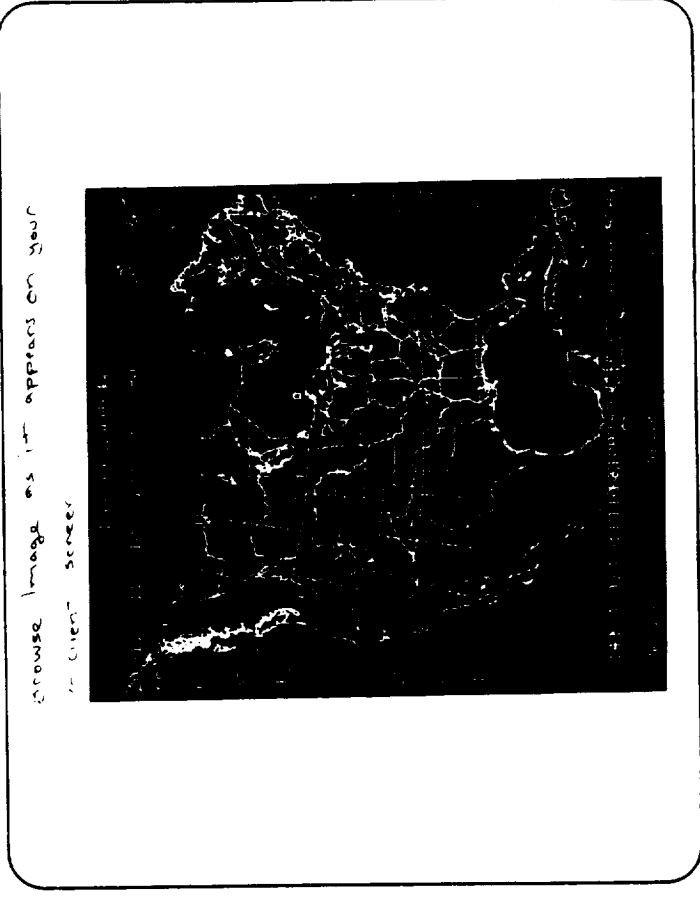
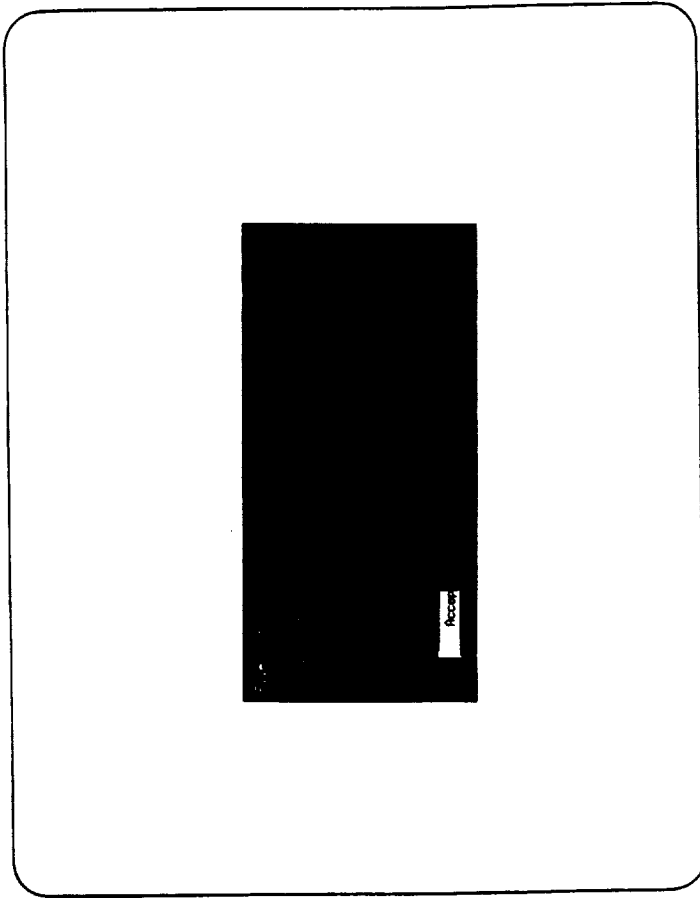


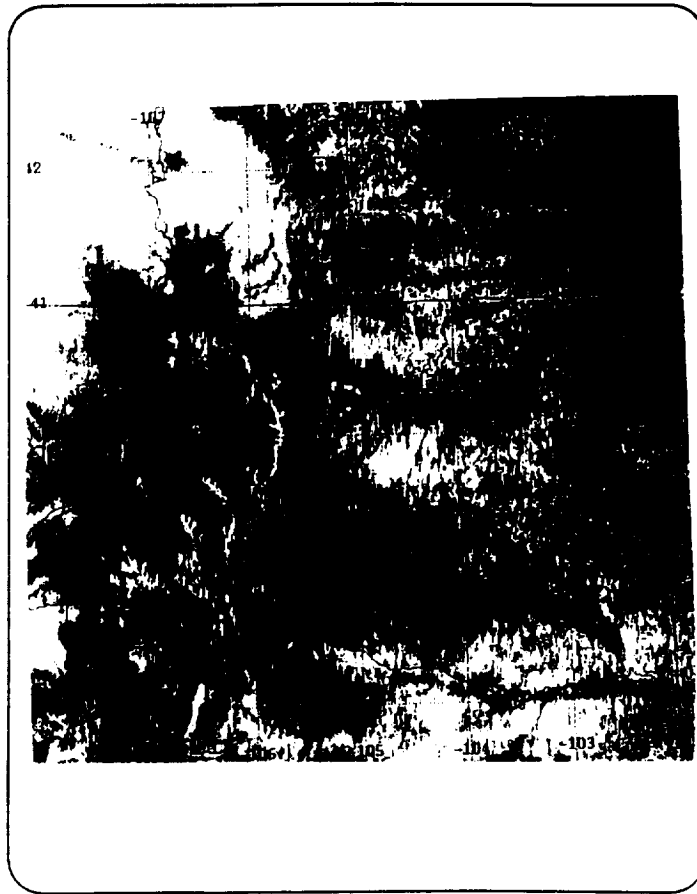
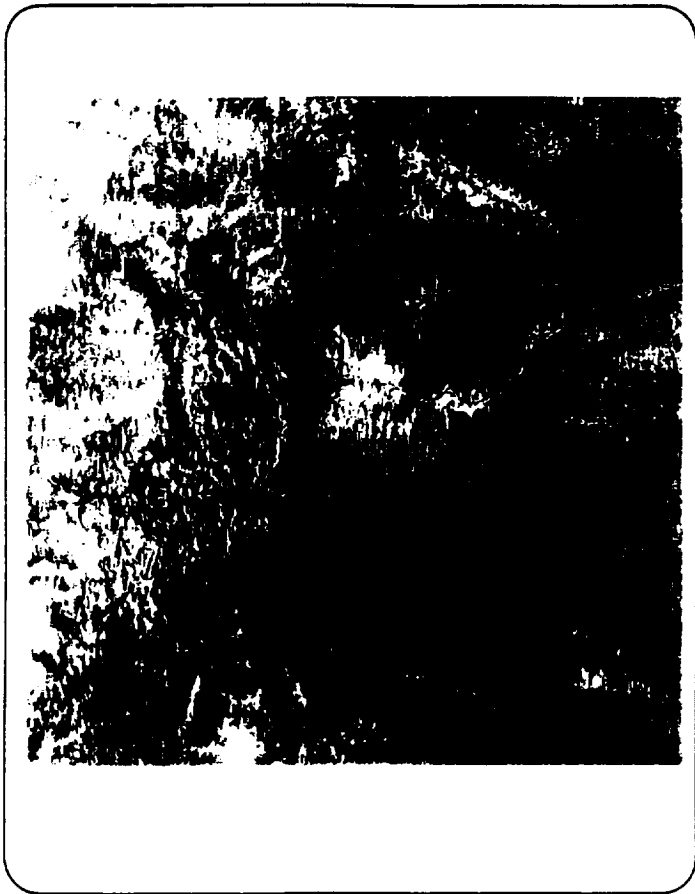
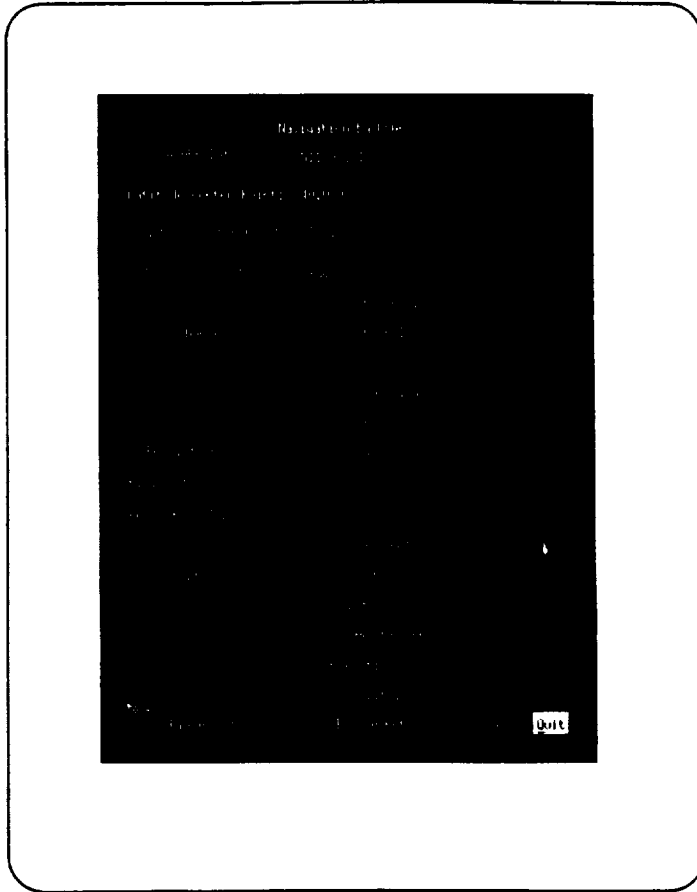
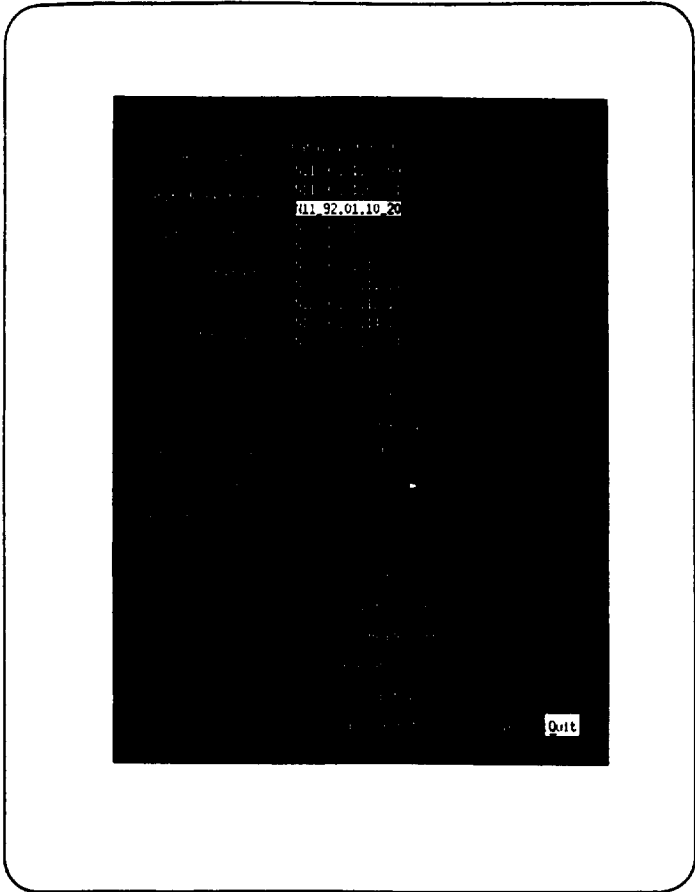
Dec 91 Apr 92 Aug 92 Dec 92 Apr 93
Months

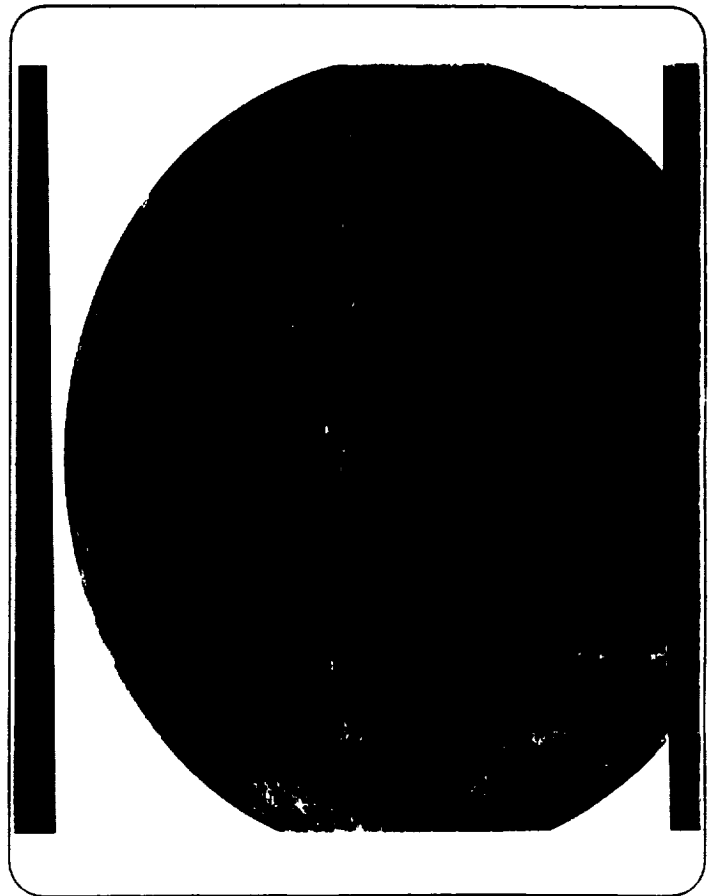
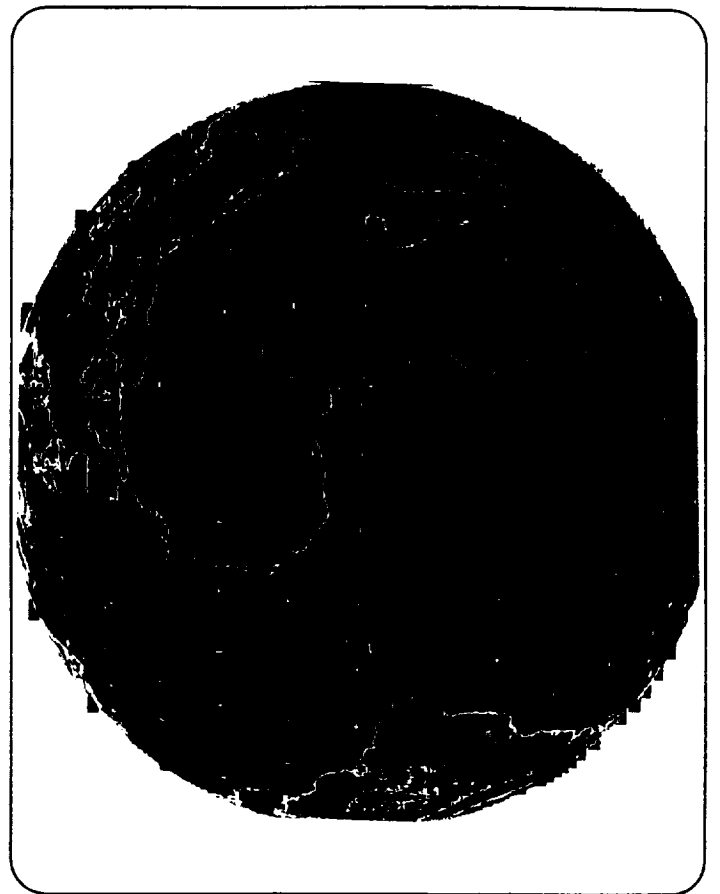
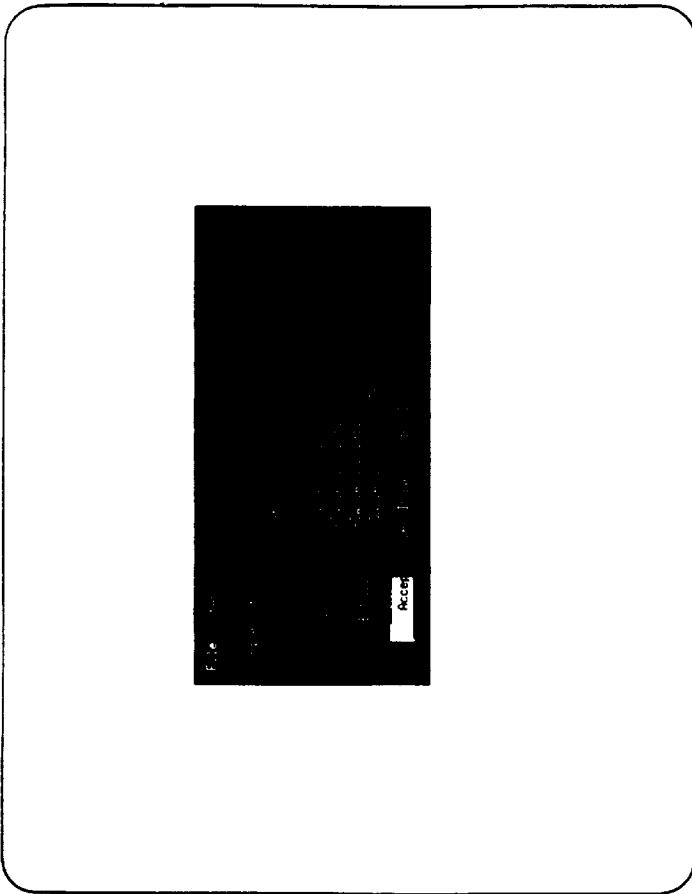


Navigated Images in Megabytes



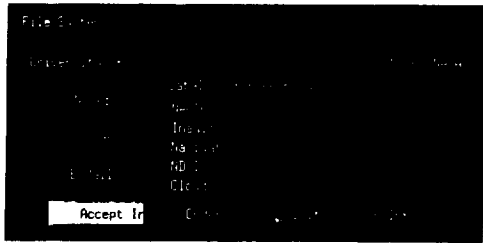




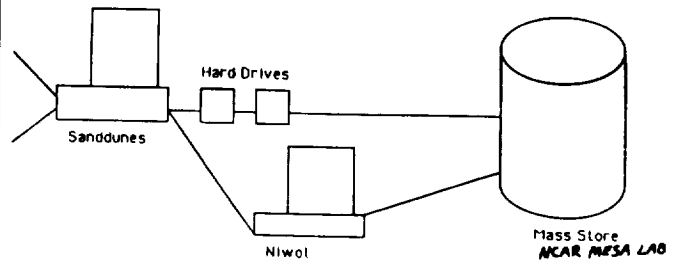


E-124

ORIGINAL PAGE IS
OF POOR QUALITY



First System Design For Images Only



*User telnets Sanddunes, and places order for Image

*Sanddunes sends order to Niwot

*Niwot places order to Mass Store

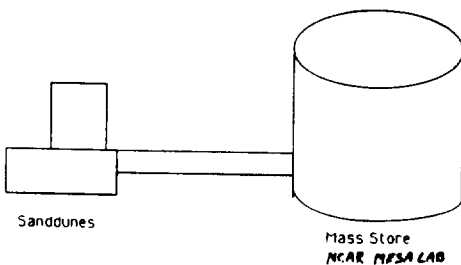
*Mass store sends sequence number to sanddunes (mail)

*Mass store sends file to sanddunes hard drives

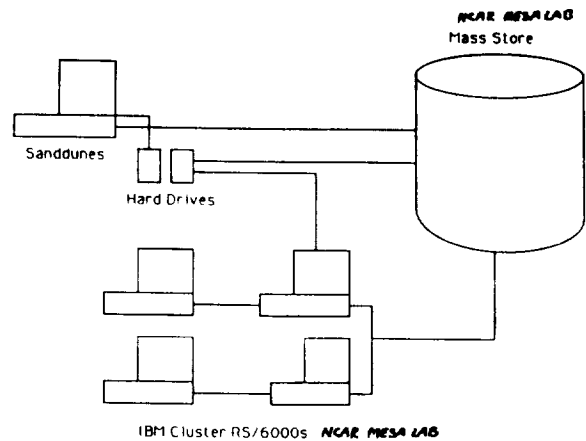
*Sanddunes parses mail for sequence number, name and file name

*Sanddunes renames file to correct name and puts file in /usr/users/ftp/usersname

Second System Design For Images Only



Navigate System Design



* User telnets to sanddunes to browse and place orders

* Sanddunes places order to Mass Store

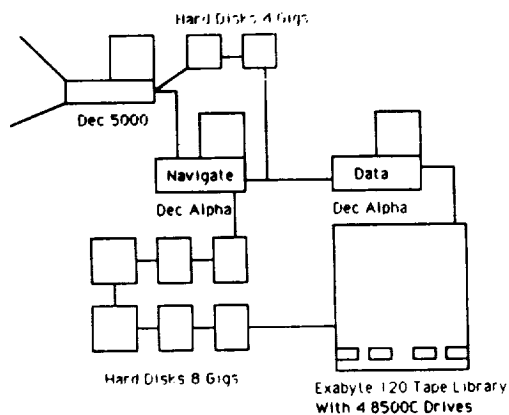
* Sanddunes spools a exec fork until file returns

* Sanddunes parses mail to rename file and find user

* Sanddunes sends navigate command to Cluster

* IBM 6000 navigates the images send message it is done

* Sanddunes moves navigated images into /usr/users/ftp/usersname



- * System designed to be totally self sufficient
- * All computers will be hard wired together for speed
- * Dec 5000 only handles browse, orders and ftp transfer
- * Dec Alpha (Navigate) only does navigations (allow spooling)
- * Dec Alpha (Data) only processes tape requests (Reads/Writes)
- * Exabyte 120 with compression drives will hold 1 terabyte of data

EOSDIS TESTBED FUTURE PLANS

- EXPAND PRESENT SYSTEM TO INCLUDE GLOBAL AVHRR DATA VIA THE CU DOMSAT ANTENNA FOR THE YEAR 1994; ON A TRIAL BASIS DUE TO THE LARGE VOLUME OF DATA
- ASSESS THE CAPABILITY OF THE NEW INDEPENDENT STORAGE SYSTEM TO HANDLE THE GREATER VOLUMES OF DOMSAT AND OTHER SATELLITE DATA
- EXPLORE THE CAPACITY OF THE NEW NETWORK CAPABILITIES TO HANDLE THE TRANSFER OF ALL STORED AND NEWLY COLLECTED DATA
- TRANSITION PRESENT SYSTEM TO V0 EOSDIS BEFORE DEC 1994; SYSTEM TO BE LOCATED EITHER AT JPL OR GSFC (OR BOTH)
- SHUT DOWN EOSDIS TESTBED DEC. 31, 1994 (OR SHORTLY THEREAFTER)

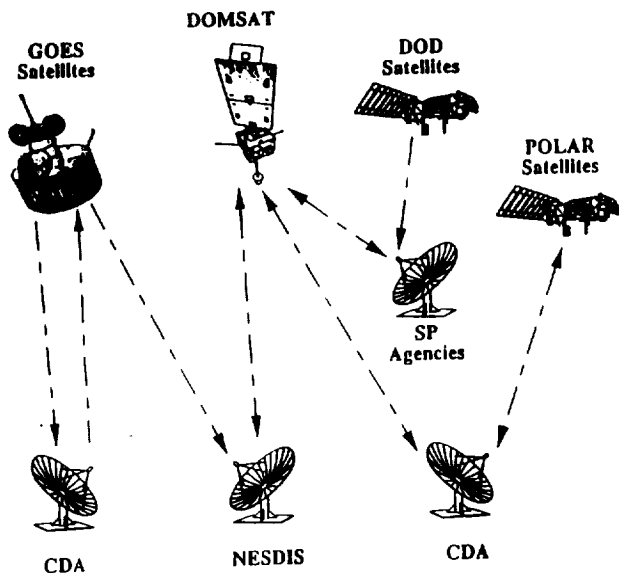


Figure 1: NESDIS Operations Context

Data Category	Data Type	Volume Per Day
GOES	AAA	3.7 GB
TIROS	GAC HRPT LAC	1.1 GB 1.7 GB 2.4 GB
SP	SSM/T SSM/I (raw) RTNEPH	56 MB 210 MB 31 MB
METEOSAT	Mapped Image	TBD

Figure 4: Sample - Estimated Current Ingest Volumes

EOSDIS TESTBED TECHNOLOGY TRANSFER

- DISPLAY AND DATA MANIPULATION SOFTWARE TRANSFERRED TO APPROXIMATELY 2,000 USERS (PLUS JPL DIST OF IMAGIC > 1,000); USERS INCLUDE EDUCATION, GOVERNMENT AND INDUSTRY
- NAVIGATION SOFTWARE HAS BEEN CUSTOM HOSTED AT OVER 8-10 SITES IN THE U.S. (ALSO RUN AT 5 DIFFERENT FOREIGN INSTS)
- CU/CCAR WILL BE FUNDED TO WORK WITH NOAA/NESDIS ACTIVE SATELLITE ARCHIVE SYSTEM TO UPDATE THEIR DATA ACCESS SYSTEM AND SOFTWARE, BASED IN LARGE PART ON THE SUCCESS OF THE EOSDIS TESTBED
- EOSDIS TESTBED SYSTEM TO BE MERGED WITH GRADS TO PROVIDE GREATER USER ACCESS TO BOTH SYSTEMS IN AN INTEGRATED FASHION
- EOSDIS TESTBED IS CONSIDERED AS AN ARCHIVE PROTOTYPE FOR EOSDIS DEVELOPMENT
- EOSDIS TESTBED SYSTEM IS TO BE INCORPORATED WITH THE K TO 12TH GRADE EDUCATIONAL PROGRAM OF THE ASPEN GLOBAL CHANGE INSTITUTE

**DEVELOPMENT OF A TOOL-SET
FOR SIMULTANEOUS, MULTI-SITE
OBSERVATIONS OF ASTRONOMICAL
OBJECTS**

**Dr. Supriya Chakrabarti
Boston University**

August 5, 1993

**Development of a
Tool-Set for Simultaneous,
Multi-Site Observations of
Astronomical Objects**

S. Chakrabarti & S.D. Godlin
Dept. of Astronomy, Boston University
(godlin@veeb.bu.edu)

J.G. Jernigan
SSC/UC Berkeley

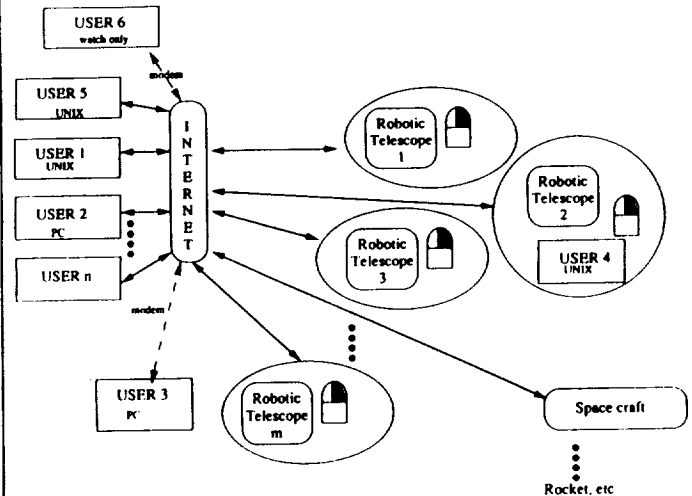
W. Coertnik
Systems Solution

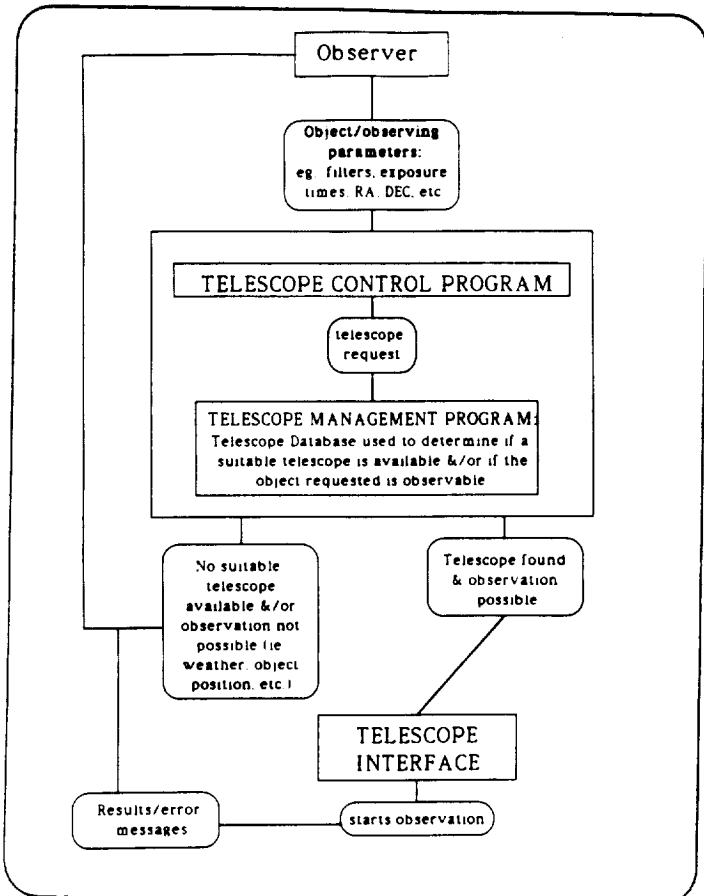
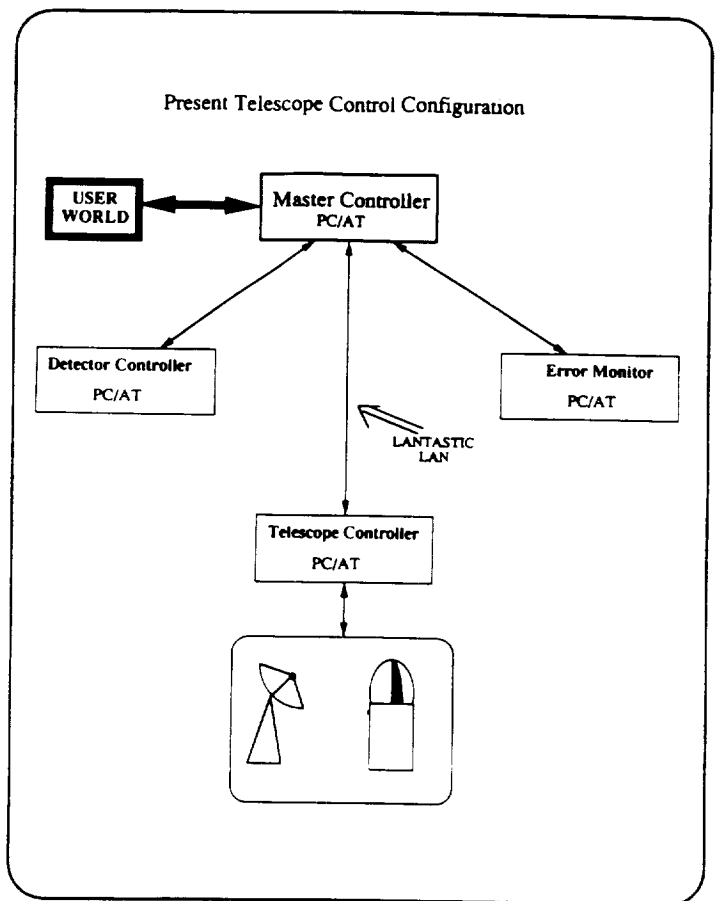
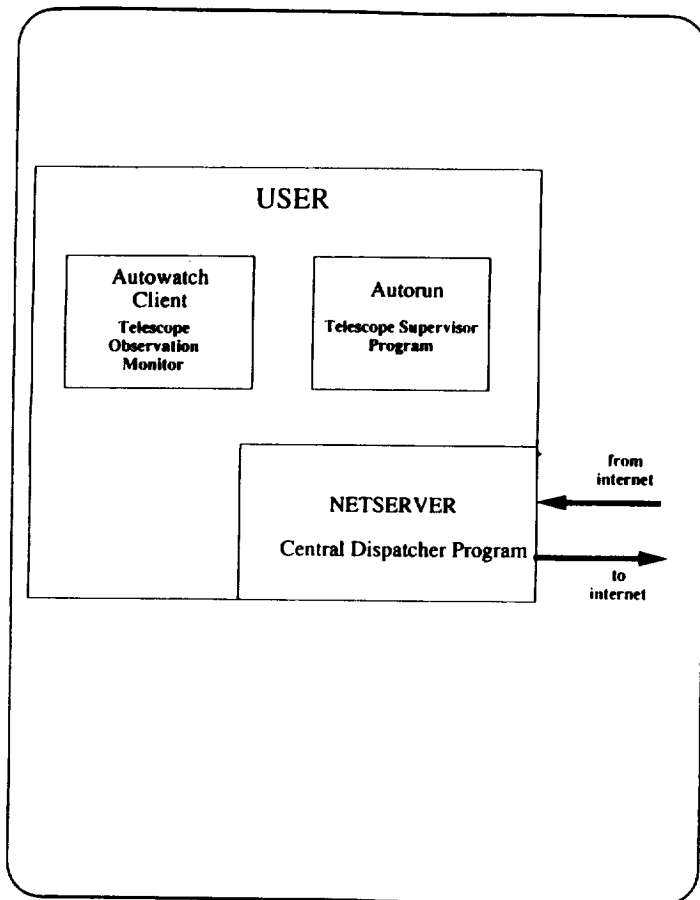
R. Genet & D. Genet
AutoScpe Corp.

Main Goal:

Make this system available to all
astronomers...amateur and
professional

GLOBAL OVERVIEW





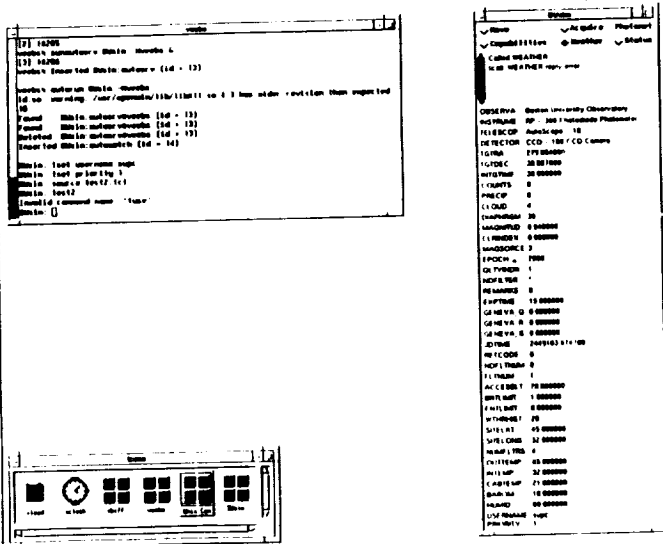
autorun.c

1. Links the user to one telescope at a time
2. Provides a graphical user interface panel which displays:
 - telescope information
 - parameter values
 - scrollbar window where text messages are displayed.

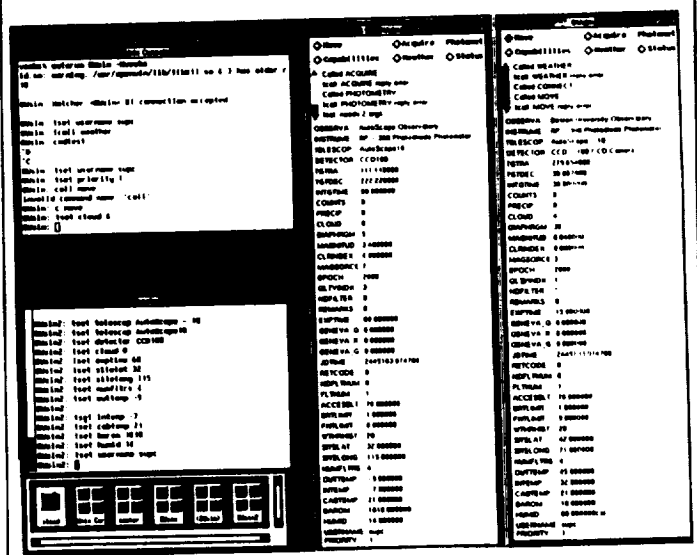
control.c

1. Allows simultaneous requests to more than one telescope from a single user
2. Each user/telescope link has a graphical user interface panel for control of each individual telescope by the observer.

Autorun alone



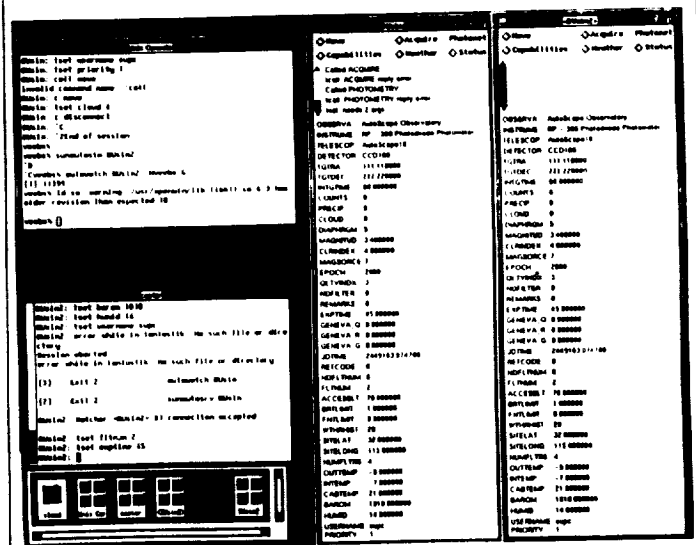
Simultaneous, Multiple Telescope



autowatch.c

1. Allows monitoring of an observation in progress
2. A user can have several 'autowatch' processes running simultaneously, each monitoring a different telescope
3. Same graphical user interface panel used in autorun.c
 - All controls in the autowatch.c panel are disabled except the scrollbar in the window where text messages are displayed. No interaction between the 'watcher' and the telescope is allowed.

Autorun in Autowatch



```

Vega - 4 LYR
SIMPLE = - - file does conform to FITS standard
BITPIX = 8 - number of bits per data pixel.
NAXIS = 0 - number of data axes
EXTEND = - - FITS dataset may contain extensions
OBSERVE = 'AutoScope Observatory'
INSTRUM = 'RP-300 Photodiode Photometer'
TELESCOP = 'AutoScope-10'
ACCSBLT = -0.000E+01 /
DETECTOR = 'CCD-100 CCD Camera' /
FNTLIMIT = 8.000E+00 /
SITELAT = 8.488E-314 /
SITELONG = 2.122E-314 /
MDFILTER = 'landolt' /
DIAPHRAG = 30 /
MAGNITUDE = 4.000E-02 /
CLRINDEX = 0.000E+00 /
MAGSORCE = 3 /
OLTYINDX = 1 /
REMARKS = 0 /
EXPTIME = 1.500E+01 photometer results - U filter
COUNTS = 0 /
GENEVA_Q = 0.000E+00 /
GENEVA_R = 0.000E+00 /
GENEVA_G = 0.000E+00 /
JDTIME = 2.4491E+06 /
FLTRMUM = 1 /
RETCODE = -1 /
EXPTIME = 2.000E+01 photometer results - B filter
COUNTS = 0 /
GENEVA_Q = 0.000E+00 /
GENEVA_R = 0.000E+00 /
GENEVA_G = 0.000E+00 /
JDTIME = 2.4491E+06 /
FLTRMUM = 2 /
RETCODE = 11 /

```

Networking Issues

- Robustness
 - No single error should crash the system
- Fault Tolerance
 - syntax and system errors
- Security
 - Password
 - User ID
 - Encryptions (not yet implemented)
- Data Compression (not yet implemented)

Software Issues

- Portability
 - UNIX, C
- Support
 - Multi-Vendor Platform
 - Sun UNIX
 - MS-DOS
- Public Domain

Networking - NETSERVER

- Central Dispatcher for entire network of telescopes and observers
 - User connected to telescope by Modem
 - PPPRELAY Procedure
 - User linked to telescope via Internet
 - autorun Procedure
 - User at telescope site

User Interface

- UNIX
- PC
 - Most amateur astronomers use PC's
- Present User Interface Software
 - Tool Command Language (tcl) in C
 - ToolKit (tk) in C
- FITS Formatted Data Files
 - FITS is the standard astronomy format
- X-windows
- Graphical User Interface
- Scheduling
 - Priority based

Users

- Located at Telescope
- Remote Observer
 - Connected via Internet or modem
- Remote 'Watcher' - viewing permission only
 - (No interaction allowed)

Telescope Interface

- UNIX
- PC
- Modes of Operation:
 - 'Batch' mode - script written in tcl
 - full or partial night observing routine
 - fully automated
 - Interactive mode
 1. graphical user interface (or)
 2. ATIS individual commands
- ATIS Batch files

Status

- Hardware Exists
- All Networking Tools in Place
- Some User Interface Written

Test Results

- **ATIS session via Internet**
 - 'Batch' & Interactive/real-time Observing
- **ATIS session using modem link**
 - 'Batch' & Interactive/real-time Observing
- **Multiple Telescope Session**
- **Used tcl scripts for specific observing sequence**
 - Scripts written by User
 - Sent to telescope by internet and modem
- **Robustness:**
 - During Multi-Telescope Session:
 1. Disconnected one telescope
 2. Connected to another telescope
 3. Used 'autowatch' on one or more sites
 - Discontinued autowatch during autorun use
- **Fault Tolerance:**
 - Syntax errors resulted in 'graceful' exit

Future

- **Improvements to User Interface**
 - put reference catalogues on line for guiding/tracking
 - move telescope according to object name
- **Test using real telescope**
- **Do real observing**
- **Add Other Capabilities:**
 - Imaging
 - Spectroscopy
 - Interferometry
- **Add Other Platforms**
- **Add Analysis Tools**
- **Add Artificial Intelligence Tools for Observation Scheduling?**

**GEOGRAPHIC INFORMATION SYSTEM FOR FUSION
& ANALYSIS OF HIGH RESOLUTION REMOTE
SENSING & GROUND TRUTH DATA**

**Mr Anthony Freeman
Jet Propulsion Laboratory**

August 5, 1993

**GEOGRAPHIC INFORMATION SYSTEM FOR FUSION
AND ANALYSIS OF HIGH RESOLUTION REMOTE SENSING
AND GROUND TRUTH DATA**

**Dr. Anthony Freeman
Jet Propulsion Laboratory**

PROGRESS REPORT

- Started out to adapt VICAR/BIS GIS to include models, different data sets (especially radar images).
- Achieved a working version.
- Poor user interface.

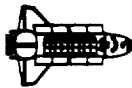
- Developed MacSigma 8.
- Supervised classification (of radar images).
- Model validation.
- Developed 3-component scattering mode.
- Developed Vegetation Map classification (MAPVEG)-Expert System.

• Overall aim—expand the user community for radar images by simplifying data display, analysis, interpretation.

- Fit 3-component scattering model to AIRSAR data.
 - Surface (BRAGG) Scatter
 - Double-bounce
 - Volume
- (Freeman and Durden, 1992)

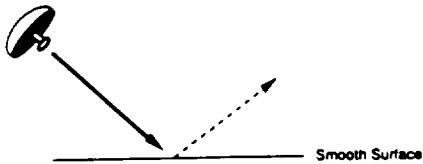
JPL

**SPACEBORNE IMAGING RADAR-C
SIR-CED**

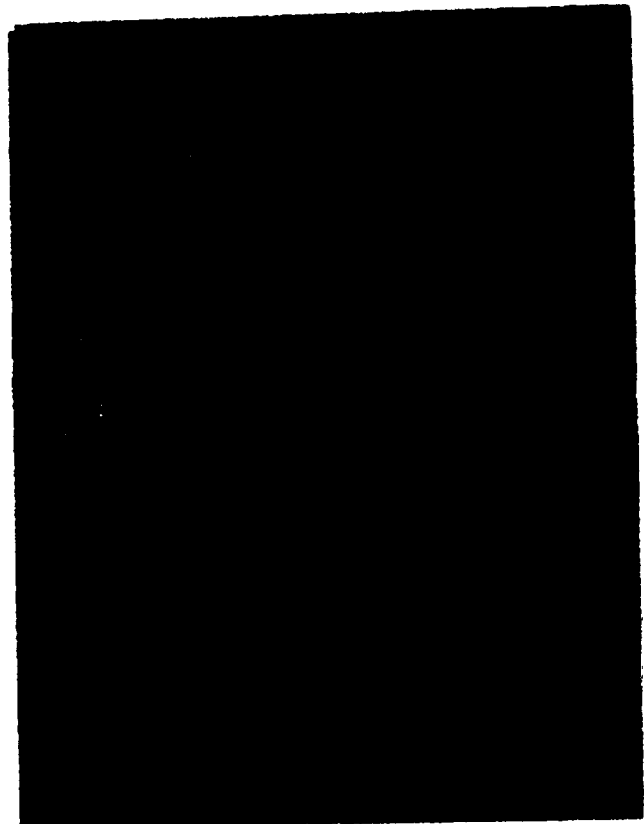
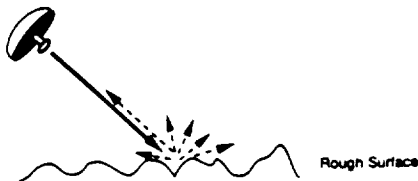


• SCATTERING MECHANISMS

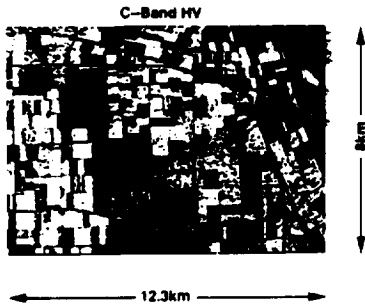
• Smooth Surface



• Rough Surface

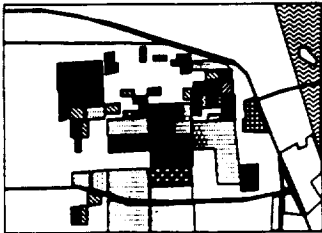


SYNTHETIC APERTURE RADAR IMAGE OF THE FLEVOLAND
AGRICULTURAL SITE IN THE NETHERLANDS



KEY:

- Water
- Forest
- Stem beans
- Potatoes
- Lucerne
- Winter wheat
- Peas
- Sugar beet
- Flax
- Bare soil
- Red beet
- Grass
- Summer barley
- Roads



•Construct an unsupervised classifier for:

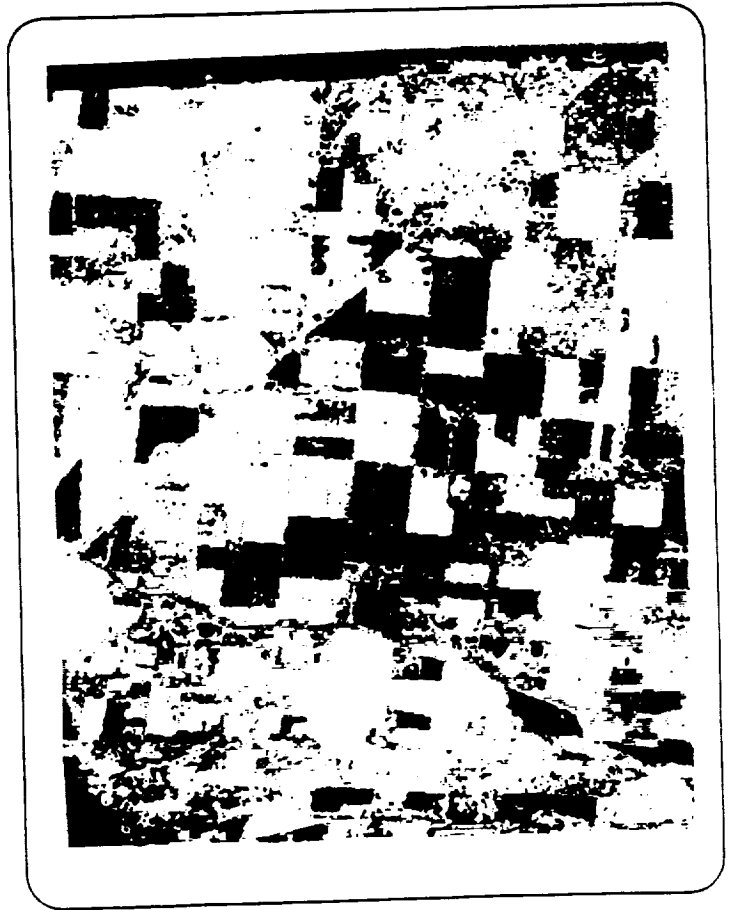
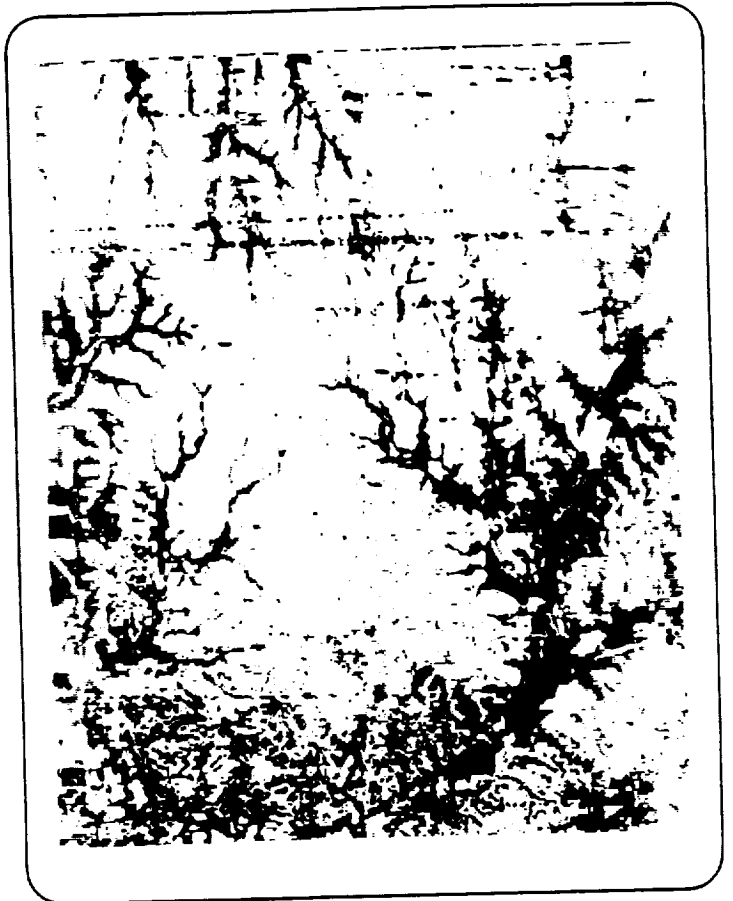
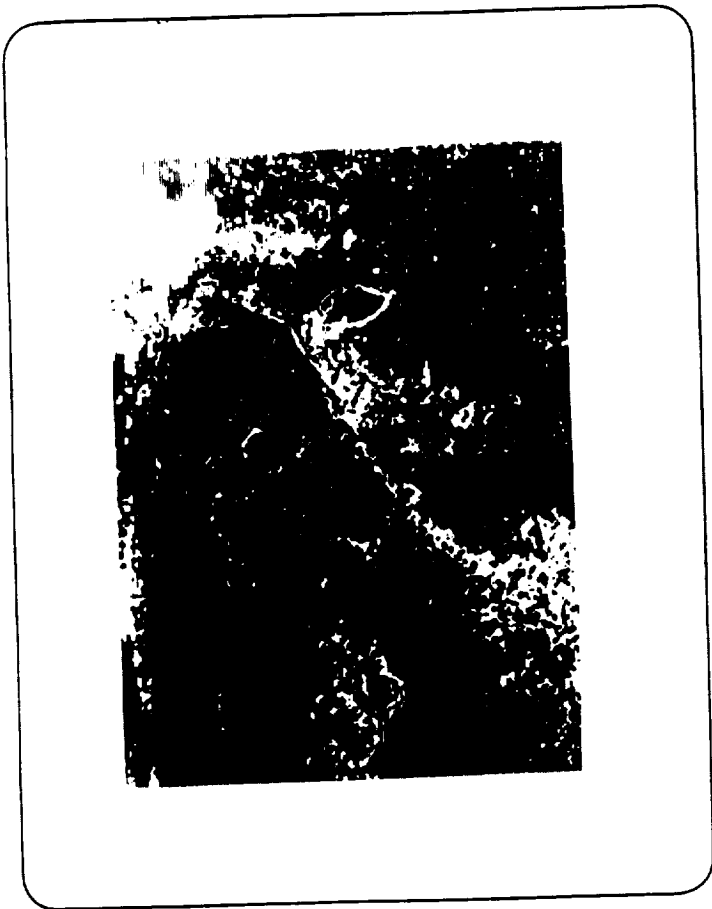
- No vegetation
- Low vegetation
- Medium vegetation
- Forest
- Urban

•Then identify vegetated areas where double-bounce is dominant.

•Classification Rules:

- No Vegetation
 - Surface scatter dominant at all 3 frequencies
 - or total backscatter very low.
- Low Vegetation
 - Volume scatter dominant at C-band but not at L-band.
- Medium Vegetation
 - L-band volume & double-bounce high.
 - P-band volume scatter low.
- Forest
 - L-band and P-band volume scatter high.
- Urban
 - Total backscatter high at L-band and P-band.
 - Double bounce > voluqe at L-band and P-band.
 - Volume scatter not dominant at C-band.
- "Double-bounce" vegetated areas:
 - Double-bounce > 1/3 volume at appropriate frequency.







-Unclassified Classification
 -Can be applied to any calibrated 3-frequency AIRSAR data.
 -Not designed for very low (< 20°) incidence angles.
 -Should be regarded as first level classification—further classes can be resolved within these simple ones.

SOFTWARE

- MacSigma 0 - released through COSMIC.
 - AIRSAR data
 - Byte data (SEASAT, SIR-B)
 - ERS-1 & ERS-1 lo-res data
 - Magellan MDR's
 - For display and analysis of radar images.
 - Export capabilities designed in.
 - MAPVEG - almost ready for release.
 - AIRSAR data
 Vegetation map
 (37 Mib) (-30Mib)
 - RAVEN - under development.
 - Display and analysis of radar images on UNIX.
 - IMAGE REGISTRATION S/W - planned
- MAC s/w to be incorporated into SIR-C Education CD.

ENVISION: A SYSTEM FOR MANAGEMENT AND DISPLAY OF LARGE DATA SETS

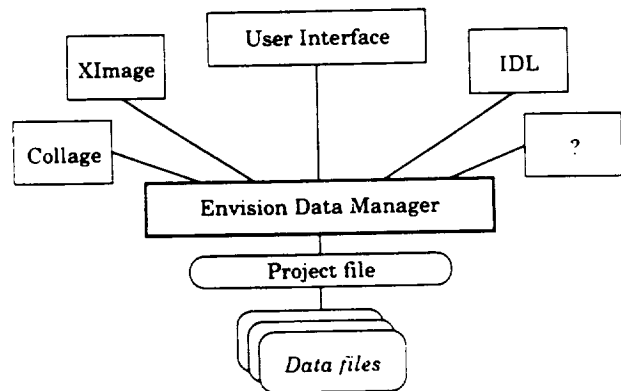
Kenneth P. Bowman
Texas A & M University

August 5, 1993

Envision: A System for Management and Display of Large Data Sets

Kenneth P. Bowman and Sridhar Pathi
Department of Meteorology, Texas A&M University

Keith Searight, John E. Walsh, and Robert B. Wilhelmson
Department of Atmospheric Sciences, Univ. of Illinois at Urbana-Champaign



Goals

- Integrate data management, manipulation, analysis, and display functions into a single interactive environment
- Use standard, portable data storage and management tools to provide access to multi-GB data sets
- Provide a simple, intuitive, collaborative, portable user interface
- Enhance the capabilities of existing interactive visualization software developed at the National Center for Supercomputing Applications (NCSA) at the University of Illinois
- Provide interfaces to other graphics display systems
- Distribute functional components
- Public domain

Envision Data Manager

Data server

Project file

- Groups one or more data files into a "Project"
- 'Virtual' dimensions
- Data 'files' larger than 2 GB
- Updating meta-data without copying netCDF files
- Adding meta-data to read-only files
- All data files remain as netCDF files (currently)

Data storage

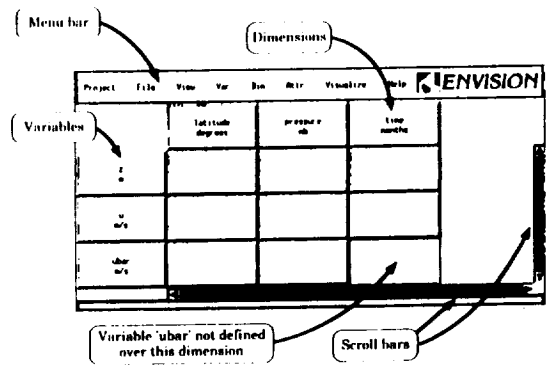
- NetCDF (UCAR/ UNIDATA)
 - Self-describing
 - Gridded, multi-dimensional arrays
 - Extensible
 - Random access to any hyperslab
 - Machine independent
- HDF (NCSA) – Hierarchical Data Format
- EOSDIS

User Interface

Table-like display of *meta*-data

- Dependent variables
- Independent variables (dimensions)
- Dimensions (sizes)
- Units or other annotation
- Can be arranged to user preferences
 - hide variables
 - hide dimensions
 - change order of variables or dimensions
- Collaborative

Envision Graphical Interface



Status and Plans

Initial release uses

NCSA XImage

NCSA Collage

IDL

Envision is available by anonymous ftp from

csrp.tamu.edu

128.194.71.50

vista.atmos.uiuc.edu

128.174.80.6

**RESEARCH AND TECHNOLOGY
ACTIVITIES AT JET PROPULSION LABORATORY**

**Dr. Amy L. Walton
Jet Propulsion Laboratory**

August 6, 1993

**Applied Information Systems Program Workshop:
Research and Technology
Activities at JPL**

A. L. Walton

August 6, 1993

Goal of Research Activities

- Provide the NASA science community with the ability to access, analyze and assimilate science data from relevant NASA and non-NASA missions, instruments and investigations.
- Partner with the science community in identifying, evaluating and developing the information systems technologies that enhance and enable science investigations.
- Provide mechanisms, technologies and organizational support for information systems in support of common instrument and science investigation needs.

Scope of Activities

Develop and exercise tools that make use of visual perception to:

- integrate and display multiple, diverse data sets
- explore and validate data sets
- track and measure features and dynamics
- provide perspective simulations
- generate scientist controlled animations

Automate and speed-up analyses of large complex data sets

Develop interactive capabilities

Extend to a heterogeneous computing environment

Incorporate techniques making the environment self-training, extensible

Speed the migration of developed tools into the science environment

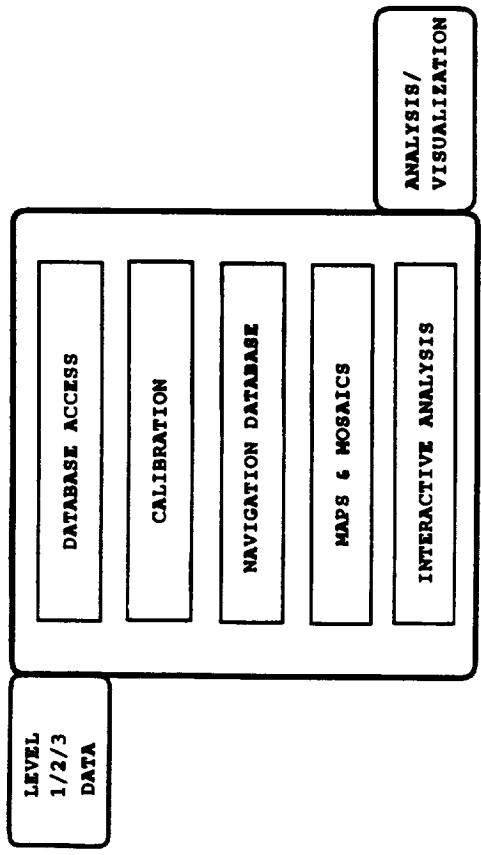
Components of the JPL RTOP Activities

- Improvements in data access:
- Data storage options
 - Ancillary data
 - Standards/Data Interoperability
- Integration into tools and services that specifically and clearly support data analysis
- Testbeds.
 - Techniques using visual perception
- Special, often limited-duration activities supported by and supporting the science community
- Data preservation
 - Newsletter
 - NRA activities

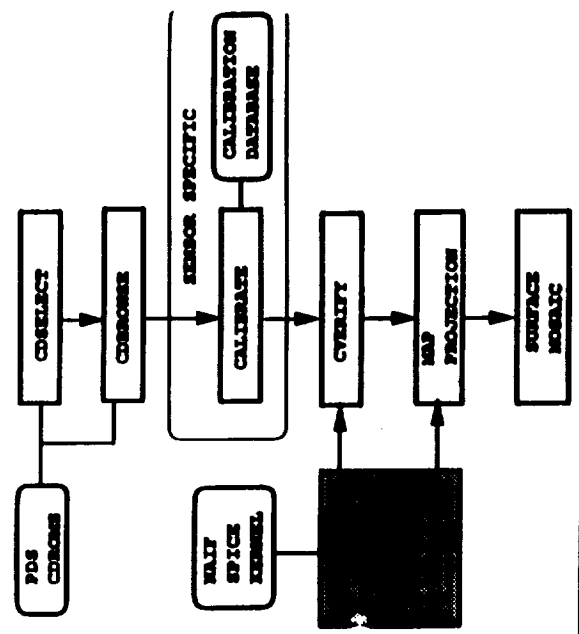
**Current RTOP Activities:
Integration and Visualization**

- IMAGING METHODS FOR MULTI-DIMENSIONAL DATA VISUALIZATION (UPN 656-65-04, D. Stanfill)**
Develop interactive visualization tools to provide scientists with the means to display and analyze three-dimensional, multi-variate, volumetric data and imagery.
- GRAPHICAL METHODS FOR SCIENCE DATA ANALYSIS (UPN 656-65-05, A. S. Jacobson)**
Provide a set of tools which permit the scientist, either alone or in collaboration, to graphically display and interpret multidimensional, multivariate data.
- SCIENTIFIC TOOLS FOR NASA/JPL IMAGE ARCHIVES (UPN 656-65-06, E. De Jong)**
Develop specific new scientific tools for use with the NASA/JPL solar system exploration archive, and develop a science analysis workstation for use with the archive.
- INTEGRATED SCIENCE ANALYSIS TESTBED (UPN 656-74-03, M. Lee)**
Demonstrate science analysis operations applied to complex remote sensing data in a heterogeneous local area network computing environment.

**PLATO
(Planetary Analysis Tools set)
BRIDGING THE GAP**

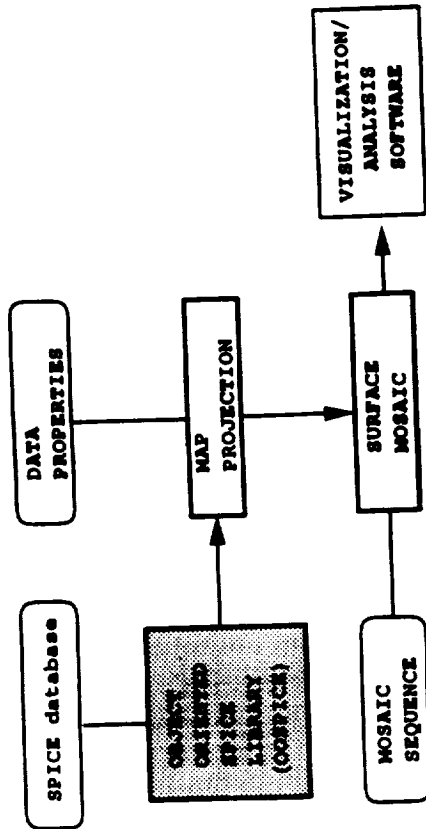


**PLATO
(Planetary Analysis Tools set)**



PLATO

MAPS and MOSAICS



Future Directions

- "Desktop Science"
 - move from timesharing/LAN to WAN
 - use smaller, local machines to access varied computing resources (supercomputers, data repositories)
 - ability to use varied data formats, multi-vendor hardware and software
- Real-time, interactive data manipulation and display of multiple, large, complex data sets.
- Environment conducive to rapid prototyping of science analysis tools (displays, controls)
 - software substrate in which user tools can easily be added (object oriented, C programming, individual tools coded as objects through a message-passing protocol)
 - application-specific tools constructed by connecting generic tools to application-specific data.
 - intuitive graphical user interface and data access, readily learned by novice.
- Scientific discoveries using visual products and workstations.
 - Video segments of all planetary systems from Mercury to Neptune

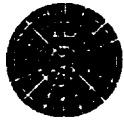
Summary

- The ability to rapidly evaluate large data volumes, complex data sets, and combinations of data sets has become a major need in the scientific research process.
- Emerging information systems tools and techniques provide a substantial opportunity to modify and extend data analysis capabilities -- in real time and in conjunction with science users.
- The creation of tools and capabilities easily used by more than one project or application has "economies of scale," and encourages the use of other elements of the Information Systems Program (computing, networking, data archives).

NAVIGATIONAL ANCILLARY INFORMATION FACILITY: AN OVERVIEW OF SPICE

Dr. Amy Walton
Ames Research Center

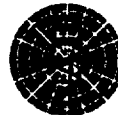
August 5, 1993



Navigation Ancillary Information Facility

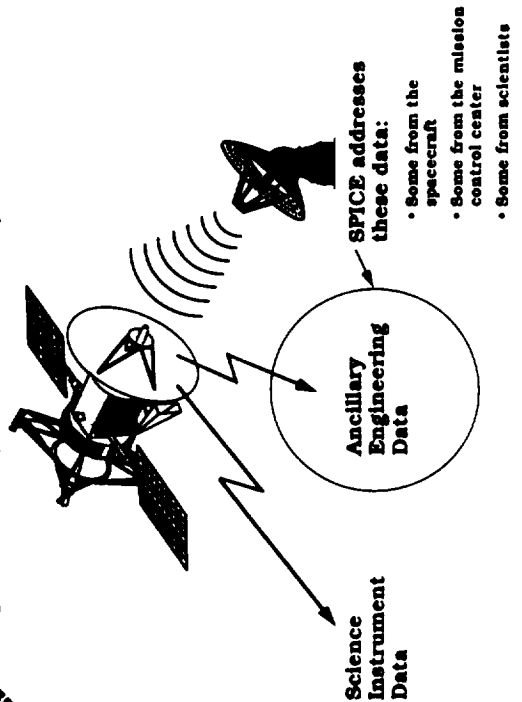
An Overview of SPICE

7/11/93



Space Science Data: Two Kinds

Navigation Ancillary Information Facility



What are Ancillary Data?

Navigation Ancillary Information Facility

- "Ancillary data" are those that tell:
 - when an instrument was taking data
 - where the spacecraft was located
 - how the spacecraft and its instruments were oriented (pointed)
 - what was the size, shape and orientation of the targets being observed

What are Ancillary Data?

Navigation Ancillary Information Facility

- "Ancillary data" can also tell you ...
 - how the instrument was acquiring data (operating mode)
 - what else of possible relevance to sensor data analysis was happening on the spacecraft or in the ground data system



SPICE System Components

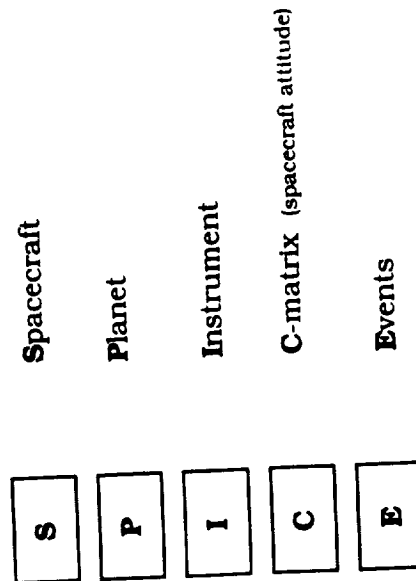
Navigation Ancillary Information Facility

The principal SPICE system components are:

- Data files
 - Software
- Also part of SPICE are:
- Standards
 - Documentation
 - User support
 - System maintenance

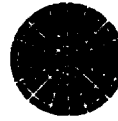
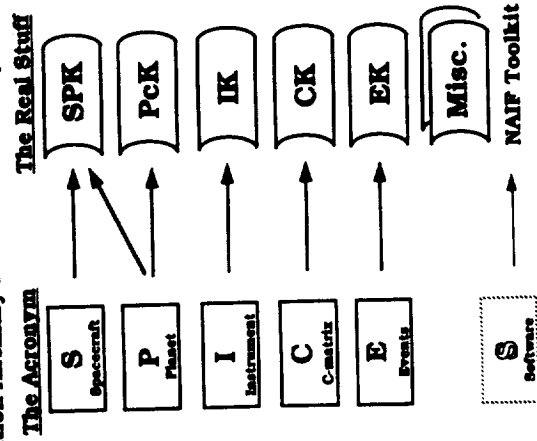
The "SPICE" Acronym

Navigation Ancillary Information Facility



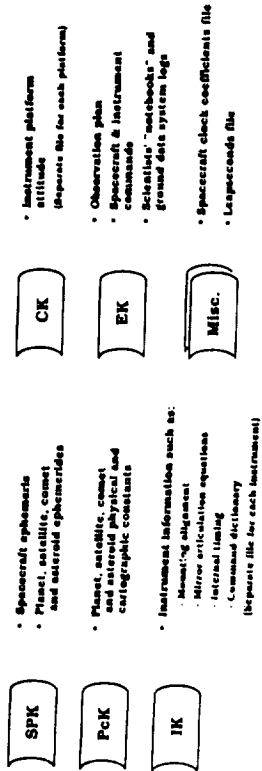
What's In the Acronym ?

Navigation Ancillary Information Facility



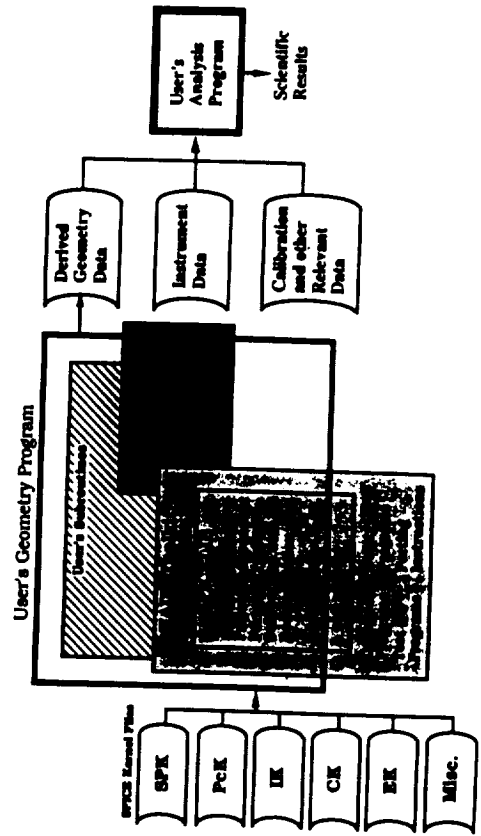
SPICE System Contents

Navigation Ancillary Information Facility



Another Possible User Scenario

Navigation Ancillary Information Facility



What's SPICE Good For ?

Navigation Ancillary Information Facility

- Space science data analysis
 - This was the original NASA motivation
- SPICE is also being used for ...
 - Science observation planning
 - From interplanetary, earth orbiting and terrestrial platforms
 - Mission evaluation, from a science perspective

Current SPICE Applications

Navigation Ancillary Information Facility

- Voyager
 - SPICE development testbed
- Magellan
 - SPK ephemeris files used in SAR and altimeter data analysis
- Galileo
 - Fully operational SPICE system for science data analysis and some observation planning and mission operations engineering support
- Mars Observer
 - MARS 94
- Clementine
 - Caseini
- Hubble Space Telescope
 - Observation planning system for solar system (moving) targets built upon SPICE components
- Other flight project and terrestrial programs
 - Ephemeris component (SPK) widely used for observation planning and data reduction



Future SPICE Applications

Navigation Ancillary Information Facility

- Examples of other possibilities for application of SPICE include the following:
 - Planetary: MESUR, NEAR, Discovery Program
 - Astrophysics: Radioastron, SPEKTR-R/G, SIRTf, AXAF
 - Earth Science: EOS, QUICK-LIS
 - Space Physics: INTERBOL
- Similar future NASA and foreign flight projects are also logical candidates for use of SPICE
- Use of SPICE as a component of a space science education program at the university level

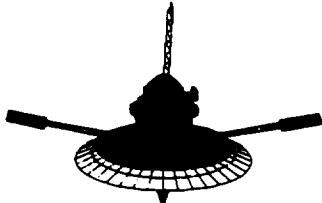
USING THE NAIF SPICE KERNEL CONCEPTS AND THE NAIF TOOLKIT SOFTWARE FOR GEOMETRY PARAMETER GENERATION AND OBSERVATION VISUALIZATION

Karen E. Simmons
University of Colorado

August 6, 1993

Using the NAIF SPICE Kernel Concepts and the NAIF Toolkit Software for
Geometry Parameter Generation and Observation Visualization

Karen E. Simmons
LASP, University of Colorado



GGGG

Galileo Geometry and Graphics Software

AISRP Workshop - Aug 3-6, 1993 - KES -



GGGG

HISTORY

Better Navigation Data Needed:

- Supplemental Experiment Data Records (SEDRs) were static products. They had conceptual flaws with respect to spacecraft, navigation, pointing and science instrument design improvements.

==> SPICE Kernels:

- Knowledge can improve
- Updates are easy
- Updates do not depend on Project
- Customization is manageable
- Archiving is simpler
- Documenting might be harder

AISRP Workshop - Aug 3-6, 1993 - KES -



GGGG

HISTORY

Kernels created the need for standardized parameter definition and generation:

- Was that SEDR "latitude" planetographic, planetocentric or planetodetic?
- What is Jupiter's System III West longitude?
- How do you convert EME1950 to J2000?

==> SPICE TOOLKIT

- Everyone uses the same "wheel" -- no more software duplication
- (Helps) Uniform understanding and use of geometry parameters

AISRP Workshop - Aug 3-6, 1993 - KES -



GGGS

What GGS Provides

Kernels provide data, the Toolkit provides tools

— GGS (Geometry and Graphics Software) provides the expertise to use both:

- to Provide geometry parameters
- to Visualize the geometric environment of a science observation
- to Maintain and Document Expertise

AISRP Workshop - Aug 3-4, 1993 - KES *



GGGS

GGGS is a Tool to Allow a Scientist to Understand the Complex Geometric Environment in which a Data Set was Obtained:

- Larger, distributed teams need observation design knowledge
- Use of archive data sets much later
- Presentation in a planetary science format
- Simple enough to run on affordable machines
- Interactive
- Capable of scientific analysis and presentation

AISRP Workshop - Aug 3-4, 1993 - KES *



GGGS

GGGS Development

One undergraduate student - Kirk Benell

- 3 work months

Basic Geometry Programs

Spherical presentation routines contributed by Doyle Hall

- 3 work months FORTRAN Graphical Interface (TEK)

GGGS FORTRAN Version

XWindows and IDL Interface

- 5 work months

GGGS Beta Version

User Inputs, Versions, Platforms, Documentation, User Servicing

GGGS Version 1.0

AISRP Workshop - Aug 3-4, 1993 - KES *



GGGS

GGGS Provides Geometry Parameters

ACTIVE Display:

- Trackball of Incidence, Emission, Phase angles on Body (or rings)
- Call-up Window with extensive parameter list for BoreSight and "corners" (choose a FOV)

HARDCOPY (or files):

- Printable table of Footprint parameters
- Graphics in selectable laser printer formats

FOOTPRINT Files:

- IDL / FORTRAN / C readable file of parameters for BoreSight and corners
- for Selectable Begin, End, and Delta time

AISRP Workshop - Aug 3-4, 1993 - KES *

SCIENCE OBSERVATION VISUALIZATION

GGGS

After-the-fact animation of the commanded sequence:

- FOVs at designated time intervals
- Selectable grids / bodies / features
- Spun section view available

Opportunity Investigation:

- Tour evaluation
- Geometry Specific opportunities: occultations, feature tracks, etc...
- Facilitates scientific intuition and creativity

Planning and Design:

- FOV placement, orientation, coverage, smear, timing and duration considerations
- Geometry constraints
- Optimization for better science return

AUSRP Workshop - Aug 3-6 1993 - KES

GGGS Maintains and Documents Expertise

GGGS

Actively - Mostly through Kernels:

- Spacecraft configuration, any scan platform configuration, operation constraints
- Instrument knowledge
- Generation history provided with kernels (usually)

Passively:

- Distributed users of same system
- Complete User, Installation, and Internals Manuals
- Several On-Line Help features
- Long term archive
- End-to-end system: design-to-observation-to-analysis
- History created automatically with Footprint files and kernels

AUSRP Workshop - Aug 3-6 1993 - KES

SCIENCE OBSERVATION VISUALIZATION

GGGS

After-the-fact animation of the commanded sequence:

- FOVs at designated time intervals
- Selectable grids / bodies / features
- Spun section view available

Opportunity Investigation:

- Tour evaluation
- Geometry Specific opportunities: occultations, feature tracks, etc...
- Facilitates scientific intuition and creativity

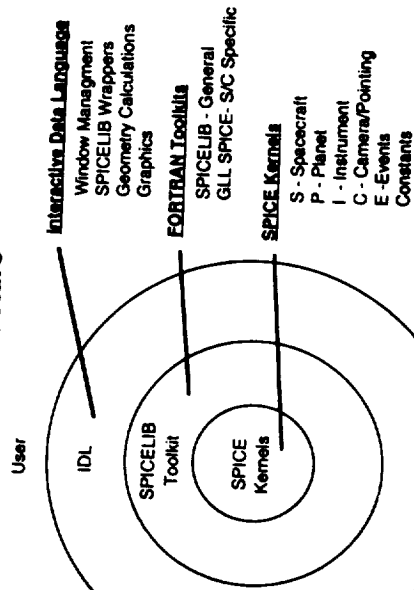
Planning and Design:

- FOV placement, orientation, coverage, smear, timing and duration considerations
- Geometry constraints
- Optimization for better science return

AUSRP Workshop - Aug 3-6 1993 - KES

GGGS Structure

GGGS



GGGS Software written by Kurt Benzel

AUSRP Workshop - Aug 3-6 1993 - KES

SPICE Kernels

GGGS

I - Kernel (Instrument)

- Mounting Location with respect to coordinate system
- Size / orientation of FOV
- Any other information, comments
- Calibration
- Filter information

E - Kernel (Event, or Experimenter's Notebook, 7or Expert)

- EVENTS - Complete sequence of events
- Experiment events - hardware events or configuration
- Engineering information: temperature, voltage, other science data
- Knowledge system - is an Expert needed?

AUSRP Workshop - Aug 3-6 1993 - KES

LESSONS LEARNED - SPICE / TOOLKIT

GGGS

Exceptionally Viable System

Kernel Generation History Documentation needs improvement - kernels are a mission product, but guidelines are needed (perhaps from PDS archive activity)

E-Kernels are still in development; they represent a vital link.

- Perhaps the most difficult to define
- Active projects will help in their development

AISRP Workshop - Aug 3-6, 1993 - KES •

LESSONS LEARNED - GGS

GGGS

Mission Independence:

- GGGG / PGGG / CGGS / VGGS ==> GGS
- A generic Geometry and Graphics Software package with "handles" to provide mission dependent aspects

Computer Platform dependence:

- Installation differences
- Data format / interaction differences
- IDL platform versions

AISRP Workshop - Aug 3-6, 1993 - KES •

LESSONS LEARNED - SPICE / TOOLKIT

GGGS

Exceptionally Viable System

Kernel Generation History Documentation needs improvement - kernels are a mission product, but guidelines are needed (perhaps from PDS archive activity)

E-Kernels are still in development; they represent a vital link.

- Perhaps the most difficult to define
- Active projects will help in their development

AISRP Workshop - Aug 3-6, 1993 - KES •

IDL

GGGS

Chose IDL for Scientific interaction with data and geometry

Added Benefits:

- Good windowing manager
- Reliable Maintenance

Concerns:

- Cost
- Upward compatibility

AISRP Workshop - Aug 3-6, 1993 - KES •

Technology Infusion

GGGS

- Students to Industry
- NAIF: SPICE Kernels and SPICELIB / Toolkit

LASP : GGS

CASSINI
MARS OBSERVER
VOYAGER (for PDS)
PHOBOS

? EOS, TIMED

Flight Project Tool

- Planning and Design
- Sequence Development
- GGS Documents describe navigation expertise

AISRP Workshop - Aug 3-6, 1993 - KES •



GGGS

What's Next ?

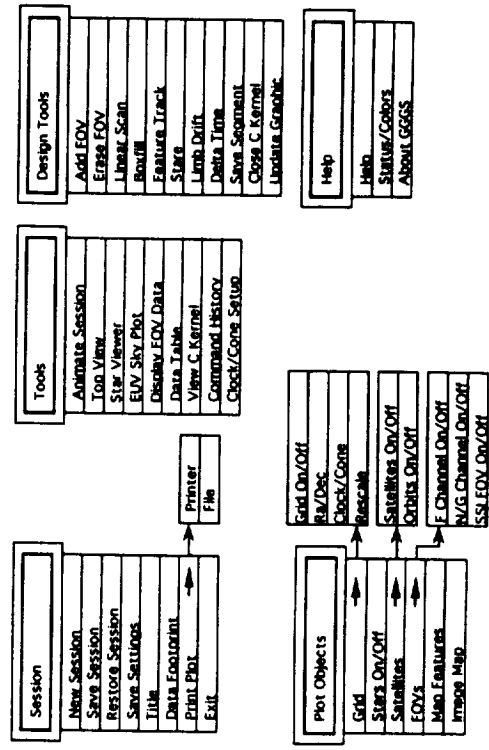
- More design-side tools - these in turn will provide better analysis tools and end-to-end tool (design-to-observation-to-analysis)
Example: Bundle of lines-of-sight for Field of View
- C-kernel Smithing with non-imaging data - methods to adjust Galileo pointing knowledge using science data; document and distribute kernels to others
- Tour Analysis Tools - enable orbit parameter tweaking to investigate the quality of science opportunities; comet and asteroid opportunities, etc...
- Investigate marriage with a Sequence Planning Tool like OASIS
- Investigate other uses: EOS, use as a PDS archive tool...

AISRP Workshop - Aug 3-6, 1993 - KES



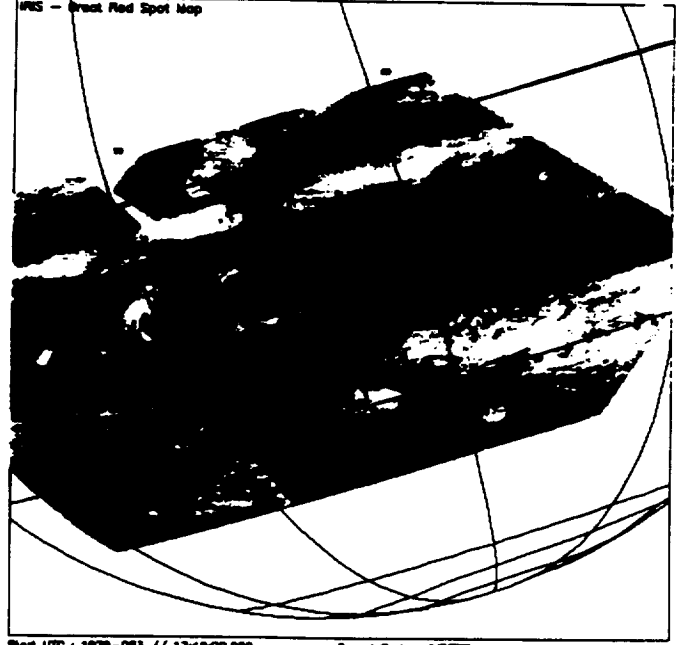
GGGS

GGGS Pulldown Menus



AISRP Workshop - Aug 3-6, 1993 - KES

VGGS VOYAGER1 IRIS FOVs with Jupiter Image



Start UTC : 1979-065 // 17:46:00.000
End UTC : 1979-065 // 18:32:00.000
Start SCLK : 2/16369:30:802
Date Time between PDV : 48.00000

Target Body : JUPITER
Target RA/Dec : 133.85/ 18.27
S/C to Body Center : 1292247. Km (16 RJ)
FOV: IRIS

AISRP Workshop - Aug 3-6, 1993 - KES

AISRP Workshop - Aug 3-6, 1993 - KES

Galileo Geometry and Graphics Software

GGGS is a software package that allows the scientist to visualize instrument viewing geometry. GGGS provides the scientist with a tool to visualize past observation geometry as well as plan future spacecraft observations. Using geometry data provided by the NAIF SPICE system and the graphics power of the Interactive Data Language (IDL), GGGS gives the scientist capabilities that had been only available to spacecraft navigation teams.

GGGS Guest Account Operation

A captive guest account has been established to provide a demo of the GGGS software package. A user can remotely log into the account and run a demo of GGGS. To run the demo, the following is required:

- An X-window display
- Access to the guest account computer (bper.colorado.edu)

Before you log in, bper must be granted access to your X-terminal device. The method in which access is granted depends on the type of system you are using. Two of these methods are:

- Unix - Enter the command 'xhost +' in the window that the guest account is logged into from.
- VMS - Select the Security option of your window manager and add bper to the access list.

When the window manager is set up, establish a connection to bper. Once a connection to bper is established, log in using the guest user name

Username: GLL_GUEST

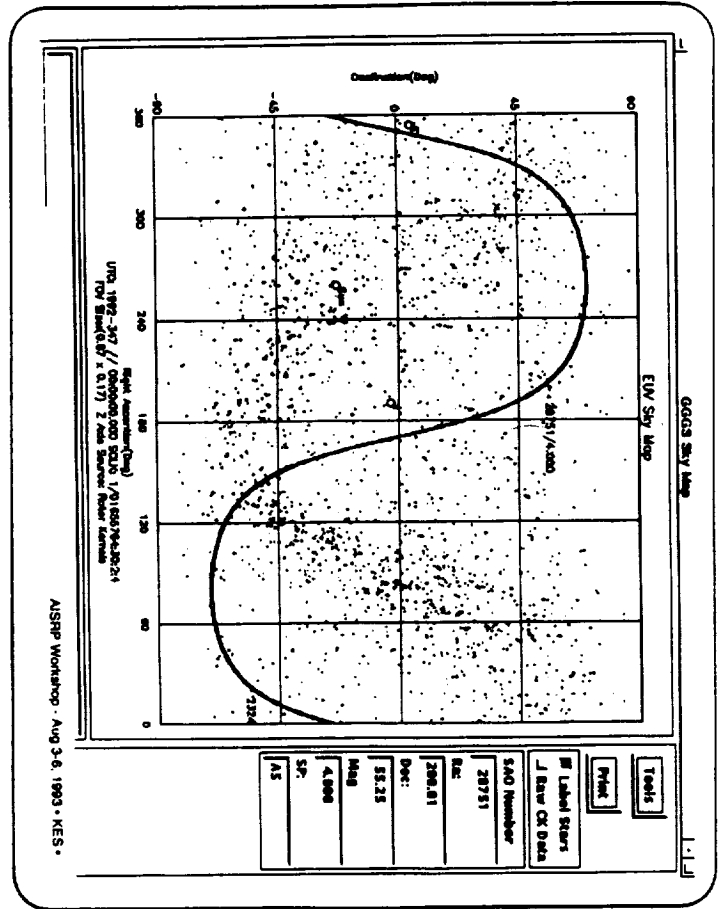
No password is required for this account. Once you have successfully logged into bper, access to your display is verified. You will be prompted for the name of your display and the transport being used (i.e. Decnet or TCP/IP). Once access is verified, a menu is printed. To run GGGS, select the "H" option.

Once the GGGS control panel appears, select the plot mode button. A data entry widget will appear. Select the Season Queue button. Select a desired observation, press done and then press Go on data entry widget.

A graphic is created. Once the graphic is drawn, try out the GGGS options. Enjoy!

Please contact Karen Simmons at or Kirk Benell for more information about GGS.

	Karen Simmons (303) 492 - 8383 SIMMONS: PISCES simmons@pces.colorado.edu	Kirk Benell (303) 492 - 5477 BENELL: PISCES benell@pces.colorado.edu
(SPAN address)		
(Internet address)		



ORIGINAL PAGE IS
OF POOR QUALITY

THE EOS/PATHFINDER INTERUSE EXPERIMENT

Dr. Michael Botts
The University of Alabama at Huntsville

August 6, 1993

THE EOS/PATHFINDER INTERUSE EXPERIMENT

Mike Botts
Earth System Science Laboratory
The University of Alabama in Huntsville
(UAH)

EOS Quarterly Review
Greenbelt, MD
May 26, 1993



UAH

What is Interuse ?

Interusability = the ease with which data sets from various sensors and disciplines can be brought together, coregistered in space and time, and correlated, analyzed, and visualized together within scientific tools.

Minimal loss or corruption of scientific information is of vital importance.

Underlying Concepts of the Interuse Experiment

A key vision of EOS is to facilitate interdisciplinary studies within earth and basic sciences = Interuse of data

Interuse is a full data path issue

Compatible decisions must be made at all points of the data path, from data set production to data use in scientific tools

Navigation plays a key role in interuse



UAH



UAH

Interuse Concerns

The present EOS direction of relying solely on previously gridded and projected Level 3 data is inadequate:

- Inevitable incompatibilities of Level 3 data
- Proliferation of redundant Level 3 data sets for different grid resolutions and projections
- Loss of intimate contact with uncorrupted Level 2 data
- Inflexibility with regard to changing gridding schemes and projections, in order to meet scientific or processing objectives



LEVEL 3

Gridding, Projection, and Navigation Issues

- (1) Like any transform, gridding and projecting data degrades data derived from sensors, through interpolation and averaging.
- (2) If two or more Level 3 data sets are not in the same projection or grid resolution, then further interpolation/averaging may need to be applied to one or more of the data sets, in order to allow correlative analysis.
- (3) Pathfinder Workshop, August 1992, illustrated difficulties in obtaining consensus on common gridding schemes. Smart tools, in addition to common grids were recommended.
- (4) For many applications, there are advantages to keeping navigation and projection later in the data path (preferably at the visualization/analysis stage).
- (5) It is possible, and advantageous, to decouple the scientific processing of data from the gridding and projection.
- (6) Advantages to "standardizing" navigation process: compatibility and ease.



LEVEL 2

Gridding, Projection, and Navigation Issues (cont)

- (7) Ability to do navigation later in path requires that:
 - (a) Level 2 data must be provided with usable navigation information
 - (b) Tools must be provided for accessing and using this information within application software
- (8) Decisions regarding navigation should be made with consideration of the full data path.



LEVEL 3

The Interuse Experiment Plan in a "Nutshell"

The Interuse Experiment will create and distribute two sets of Level 2 data to be used for testing Interuse.

Several issues must be resolved before an easily used Level 2 data set can be produced and distributed

Appropriate navigation information, and tools for navigating, gridding, projecting, generating mosaics, analyzing, and visualizing the data will accompany the data sets

Issues regarding the Interuse of Level 2 and Level 3 data will be considered



LEVEL 2

Issues Needing to be Resolved

- Generic navigation for EOS data
- Generic gridding and projection routines
- Incorporation of navigation information into HDF
- Data subsetting and compression
- Availability and compatibility of appropriate analysis and visualization tools



JPL

Data Set Scenarios

- Scenario 1: Data sets for a limited regional area over a long period of time
- Scenario 2: Data sets of global extent over a limited period of time
- Interuse Tiger Team will provide input on the areas and times desired



JPL

Key Pieces

- SPICE - generic navigation system developed at JPL for planetary missions; provides navigation data as a set of 5 SPICE kernels (Chuck Acton - JPL)
- SPICElib & OoSPICE - FORTRAN and C libraries which allow access and use of SPICE kernel information from within application software (Meemong Lee/Richard Weidner - JPL)
- PLATO - application software which accesses SPICE kernels in PDS files and allows data browse, projection, gridding, and mosaic creation (Meemong Lee/Richard Weidner - JPL)
- LinkWinds, IDL/PV-Wave, AVS Explorer, Khoros - Visualization/Analysis software which will allow scientific correlation and analysis between data sets



JPL

Interuse Teams

- Interuse Tiger Team - consists of Mike Boffs/JAH team, Chuck Acton/Bill Taber, Meemong Lee/Richard Weidner + Martha Maiden + Amy Walton meets quarterly to work out "dirty details"
- Interuse Core Working Group - consists of scientists and data producers picked by Martha Maiden meets every 2-3 quarters to provide suggestions and act as "tire kickers" for the Interuse Core Working Group ideas and progress



JPL

Phase I

Establish the Intense Core Working Group and Tiger Team for the intense experimentation

Gain familiarity with candidate data sets and established data processing techniques - generate SPICE kernels for test data

Test SPICE or other navigation systems for general use with Pathfinder and EOS data sets

Work with ESD/ISP, NCSA, and ECS contractor to incorporate navigational information and other metadata into HDF; develop a "common" HDF header for navigational data and metadata



INTENSE CORE WORKING GROUP

Phase I (cont)

Investigate capabilities for intense use of disparate Level 3 data sets

Test capabilities for browse, navigation, gridding, projection, and mosaic generation using navigational library within visualization tools

Review results of studies re: compression and subsetting capabilities for large files

Collect, modify, or develop operational tools and libraries for navigation, general gridding, projections, and mosaic production



INTENSE CORE WORKING GROUP

Phase II: Data Set Generation, Distribution, and Testing

Create Level 2 datasets using generalized navigation capabilities and place in HDF format

Store Experimental Level 2 datasets

Distribute datasets and toolkits for intense testing by SCF scientists

Provide user support for the distributed data sets and tools

Measure success of intense experimentation and establish recommendations for improving intense use of EOS data



INTENSE CORE WORKING GROUP

Phase III: Follow-up

Refine libraries and programs for general use by EOS science and data production communities

Refine and redistribute data sets, if desired or required

Work with NASA Headquarters, ESD/IS, ECS, and general EOS community to incorporate successful concepts into the EOS project

Work with software developers to encourage incorporation of navigation, gridding, and projection capabilities into analysis and visualization software



INTENSE CORE WORKING GROUP

OVERVIEW OF AMES RESEARCH CENTER: ADVANCED NETWORK APPLICATIONS

Mr. John Yin
Ames Research Center

August 6, 1993

Overview
of
Ames Research Center

Advanced Network Applications

John Yin
Head, ANA Section
Wide Area Networks and Services Branch
Information and Communications Systems Division



John Yin, NASA Ames Research Center, MS 235-111, Moffett Field, CA 94035-1000

Advanced Network Applications Overview

The Advanced Network Applications (ANA) Section at ARC provides network based solutions for NASA and other Federal Agencies through the development, implementation, and operations coordination and support of extensible high technology network services and applications.

- Network Applications and Information Center
- Network Services Development Group
- InterCenter Council for Computer Networking/Science
- Electronic Messaging Implementators Group
- InterAgency Coordinating Groups



John Yin, NASA Ames Research Center, MS 235-111, Moffett Field, CA 94035-1000

Network Applications and Information Center

The NAIC provides NASA-wide operations support of advanced network services for the NASA science and research community.

- HELP DESK providing direct customer support
- Coordination with other network information centers
- Support design & establishment of Center NJCS
- Training on new applications for users and support staff
- Appropriate documentation and users guides
- Demonstration of new applications and services to users



John Yin, NASA Ames Research Center, MS 235-111, Moffett Field, CA 94035-1000

Network Services Development Group

The NSDG develops, implements, and transitions to the NAIC state-of-the-art standards-based network applications and services. These new and advanced applications and services are based on emerging national and international standards which are extensible and useable by other Agencies.

- Interoperable X.400/SMTP electronic mail gateways
- Fully distributed X.500 directory service
- Electronic signatures for authentication of documents
- Privacy Enhanced Mail
- Packet Video to the desktop
- Wireless networks



John Y. A. NASA Ames Research Center, MS 2311, 18, Moffett Field, CA 94035-1000

Development Activities

X.400/SMTP electronic mail gateways

Provides an Agency Wide electronic messaging infrastructure which allows for completely transparent document exchanges between all host and LAN based mailers. This activity has been endorsed by HQ Code J and is coordinated through the Electronic Messaging Implementors Group (EMIG), in coordination with the ICCN.

Distributed X.500 directory service

Allows for the management and distribution of an electronic locator service common to all NASA sites. Enhances electronic mail service by supporting simplified electronic mail addresses based on user names and sites. Adopted by all major NASA Centers and ANA is currently providing support to USPS, DOD, and IRS under a proposed MOU.



John Y. A. NASA Ames Research Center, MS 2311, 18, Moffett Field, CA 94035-1000

Development Activities

Electronic Signatures and Certificate Management

Joint activity with HQ Code B, NIST, and USPS in developing and demonstrating the generation, certification, management, and use of electronic signatures for authenticating electronic documents.

Privacy Enhanced Mail

Provides a standards based end to end encryption of electronic messages. Currently supports unix workstations. Working with software vendors to support PCs and Macs.



John Y. A. NASA Ames Research Center, MS 2311, 18, Moffett Field, CA 94035-1000

Development Activities

Packet Video Development and Deployment

Project sponsored by HQ/Office of Aeronautics HPCC program for providing interoperable color packet video to unix workstations, PCs, and Macs by end of FY94 with initial demonstration on Sun workstations at NASA HQ and ARC by end of FY93.

Wireless Network Development

Demonstrate high speed (11Mbps+) wireless connectivity to portable computers with special emphasis on support for advanced applications like packet video.



John Y. A. NASA Ames Research Center, MS 2311, 18, Moffett Field, CA 94035-1000

