NASA-CR-193096

Applied Information Systems Research Program (AISRP)

Workshop II

Meeting Proceedings

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

August 11-13, 1992

Proceedings Issued By: Information Systems Branch
Flight Systems Division
Office of Space Science and Applications
NASA Headquarters

Applied Information Systems Research Program (AISRP)

Workshop II

Meeting Proceedings

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

August 11-13, 1992

Proceedings Issued By: Information Systems Branch
Flight Systems Division
Office of Space Science and Applications
NASA Headquarters

APPLIED INFORMATION SYSTEMS RESEARCH PROGRAM (AISRP)
WORKSHOP II

TABLE OF CONTENTS

SCIENTIFIC VISUALIZATION SESSION

PROGRAMMATIC ISSUES AND ANSWERS SESSION

APPENDICES

OMIT TO
PD-1

EXECUTIVE SUMMARY

The University of Colorado's Laboratory for Atmosphere and Space Physics in Boulder, Colorado was the site of the second workshop of the Applied Information Systems Research Program (AISRP). The workshop was sponsored by the Information Systems Branch of NASA's Office of Space Science and Applications (OSSA). The purpose of this year's workshop was to evaluate progress and discern topics for new AISRP research to address OSSA science division needs. To this end, AISRP investigators presented their progress to date and future direction. OSSA Information Systems Management Board science division members or their representatives participated in a panel discussion of the scientific needs and problem areas in the process of doing their discipline's science.

The Earth and space science participants were able to see where the current research can be applied in their disciplines and computer science participants could see potential areas for future application of computer and information systems research. The Earth and Space Science research proposals for the High Performance Computing and Communications (HPCC) program were under evaluation. Therefore, this effort was not discussed at the AISRP Workshop. OSSA's other high priority area in computer science is scientific visualization, with the entire second day of the workshop devoted to it.

Many of the AISRP investigations are intended for multi-disciplinary application to the Earth and space sciences. Therefore, the presentations were organized into sessions according to computer science disciplines. Joe Bredekamp of NASA/OSSA, and the program's sponsor, opened the workshop with an overall context for the program. Dave Thompson of the Ames Research Center chaired the Tuesday morning session on artificial intelligence and related areas of expert systems and neural networks. Randal Davis, of the University of Colorado and our host, chaired the Tuesday afternoon session devoted to data related research, including data compression, data archiving, data access, and data analysis. The entire day on Wednesday was devoted to the scientific visualization session chaired by Mike Botts of the University of Alabama in Huntsville. Glenn Mucklow, OSSA's Program Manager for Information Systems Research and Technology for Joe Bredekamp, chaired the workshop and Thursday's session addressing programmatic issues. Thursday included the science discipline panel and a panel on closer collaboration and cooperation between the Office of Aeronautics and Space Technology (OAST) and OSSA research and development activities with application in the Earth and space sciences. The closing discussions included plans for new research announcements in FY 1993 from the AISRP and the Center of Excellence in Space Data and Information Sciences.

The workshop addressed issues raised at last year's workshop in technology transfer across disciplines and to the broader scientific community. Action was taken to establish a software support laboratory at the University of Colorado under the direction of Randal Davis. This group will provide a minimalist capability for testbeds and demonstrations, software tool distribution, capturing user experience, and the development of data test suites. In addition, progress in other areas identified at the last meeting were covered. The progress reported in the area of data formats resulted in a splinter group discussion as did the areas of data compression and future plans for the Internet. Agreements on collaboration were made between Dr. Hansen and Dr. Jacobson, and Dr. Emery and Dr. Kinter. OAST agreed to provide some funding for AI investigators to work with the OSSA PI's.

Glenn H. Mucklow
Program Manager
Information Systems Research & Technology

51-82
ABS. ONLY

N93-25793

p. 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.

53-82
ABS. ONLY
159832
P. 1

N93-25794

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.

S3-89
ABS ONLY
159833

N93-25795

P-1

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

Principal Investigator: Dr. Supriya Chakrabarti
University of California

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.

54-82
159834

N93-25790

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.

55-82
ABS. ONLY
159835
P-1

N93-25797

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.

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

Principal Investigator: Professor Stanislav Djorgovski
California Institute of Technology

---BIBLIOGRAPHY---

Papers based in part on the work performed under the contract NAS5-31348 so far:

1. Applications of the Multivariate Statistical Analysis Package:

"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.

"Multivariate Analysis of Globular Cluster Systems in Early-Type Galaxies", Santiago, B.X., and Djorgovski, S. 1992, *M.N.R.A.S.* in press.

"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.* in press.

"Dynamical Evolution Effects on the Hot Stellar Populations in Globular Clusters", Djorgovski, S., and Piotto, G. 1992, submitted to the *Astron. J.*

"Towards the Solution of the Second Parameter Problem: the Effects of Cluster Density and Concentration the Horizontal Branch Morphology", Fusi Pecci, F., Ferraro, F., Bellazzini, M., Djorgovski, S., Piotto, G., and Buonanno, R. 1992, submitted to the *Astron. J.*

"What Determines the Stellar Mass Functions in Globular Clusters?", Djorgovski, S., Piotto, G., and Capaccioli, M. 1992, submitted to the *Astron. J.*

2. Development of the Automatic Classification and Sky Survey Analysis Tools:

"Towards a Digitized Second Palomar Sky Survey: Initial Reduction and Star/Galaxy Classification", Weir, N., Djorgovski, S., Fayyad, U., and Doyle, R. 1991, *Bull. Am. Astron. Soc.* **23**, 1434.

"Applying Machine Learning Classification Techniques to Automate Sky Object Cataloguing" Fayyad, U., Doyle, R., Weir, N., and Djorgovski, S. 1992, to appear in: *Proceedings of the International Space Year Conference on Earth & Space Science Information Systems, Pasadena, CA, February 1992.*

"Automating Sky Object Classification in Astronomical Survey Images", Fayyad, U., Doyle, R., Weir, N., and Djorgovski, S. 1992, to appear in: *Proceedings of the Machine Discovery Workshop, Ninth International Conference on Machine Learning, Aberdeen, Scotland, July 1992.*

"The Palomar Observatory - STScI Digital Sky Survey: I. Program Definition and Status", Djorgovski, S., Lasker, B., Weir, N., Postman, M., Reid, I.N., and Laidler, V. 1992, *Bull. Am. Astron. Soc.* **24**, 750.

"The Palomar Observatory - STScI Digital Sky Survey: II. The Scanning Proces", Lasker, B., Djorgovski, S., Postman, M., Laidler, V., Weir, N., Reid, I.N., and Sturch, C. 1992, *Bull. Am. Astron. Soc.* **24**, 741.

"An Analysis of the Palomar Observatory - STScI Digital Sky Survey: Catalog Construction and Initial Results", Weir, N., Djorgovski, S., Fayyad, U., and Doyle, R. 1992, *Bull. Am. Astron. Soc.* **24**, 741.

Additional papers are now in preparation.

56-82

N93-25798

159836

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

p. 2

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

N93-25799

p-2

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.

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

---BIBLIOGRAPHY---

Richard M. Keller and Michal Rimon, "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 Modelling", 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.

S8-82

N93-25800

159838

System of Experts for Intelligent Data Management (SEIDAM)

P-1

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.

59-82

ABS ONLY

159839

N93-25801

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.

510-82
ABS. ONLY
159840
N93-25802

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.

S11-82
ABS. ONLY
159341
p-1

N93-25803

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.

512-82

N93-25804

ABS ONLY 159842

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
ABS ONLY
159843
p-1

N93-25805

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.

5/4-82
N93-25806
159844
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.

515-82
159845
p-8

N93-25807

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 throughout, 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 photofractive 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.

516-82
159846

N93-25808

P-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.

---BIBLIOGRAPHY---

REFERENCES

Yen, F. (1992). "Draco - A Data Reduction Expert Assistant", in the Proceedings of the AAAI Fall Symposium on Intelligent Scientific Computation, Boston.

Miller, G. and F. Yen (1992). "The Data Reduction Expert Assistant" in the Proceedings of astronomy from large Databases II Hagenau, France, ESO.

This last reference is an invited talk.

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

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

---BIBLIOGRAPHY---

REFERENCES

1. Hua-Kuang Liu, "Self amplified optical pattern-recognition technique", Applied Optics, **31:14**, 2568-2575 (1992).

2. Hua-Kuang Liu, Jacob Barhen, and Nabil H. Farhat, "Optical implementation of terminal attractor-based associative memory", Applied Optics, **31:23**, 4631-4644 (1992).

3. S. Zhou, P. A. Yeh, and H. K. Liu, "Dynamic self-amplified photorefractive optical beam-array generation", Presented at the SPIE Annual Meeting, San Diego, July 19,1992.

4. H.K. Lui and S. Zhou, "Complex reconfigurable free-space optical interconnection via phase CGH in spacial light modulators", Presented at the SPIE Annual Meeting, San Diego, July 19,1992.

5. H. K. Liu, "Bifurcating optical pattern recognition in photorefractive crystals", Presented at the SPIE Annual Meeting, San Diego, July 19,1992. Accepted for Publication in Optics Letters.

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.

S18 82
ABS ONLY
159848

N93-25810

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

P-1

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.

N93-25812
ABS. ONLY
159849
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.

530 82
ABS ONLY
159850

N93-25812

SAVS: A Space Analysis and Visualization System

P-1

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.

S21-82
N93-25813
159851
P-1

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

Principal Investigator: Charles H. Vermillion
Goddard Space Flight Center

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.

522 32
ABS. ONLY
159852
P-1

N93-25814

A Distributed Analysis and Visualization System for Model and Observational Data

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.

APPENDIX A:

AISRP Workshop Agenda

Tuesday, August 11, 1992

Mr. Glenn Mucklow and Mr. Joseph Bredekamp from the Information Systems Branch of OSSA welcomed the meeting participants to the second AISRP Workshop. Mr. Mucklow introduced the Workshop host, Dr. Randal Davis from the Laboratory for Atmospheric and Space Physics (LASP), and representatives from OAST, CESDIS, and the Ames Research Center (ARC). Mr. Mucklow briefly reviewed the agenda and noted the schedule for the workstation demonstrations during breaks and lunch. The Workshop Agenda and Demonstration Schedule are included in this Report in Appendices A and C.

S1-82

ARTIFICIAL INTELLIGENCE/EXPERT SYSTEMS/NEURAL NETWORKS SESSION

The Chairman of first session, Dr. David Thompson from ARC, introduced the presentations on Artificial Intelligence, Expert Systems, and Neural Networks.

<u>Multivariate Statistical Analysis Software Technologies for Astrophysical Research Involving Large Data Bases</u>

Dr. S. G. Djorgovski
California Institute of Technology

The Principal Investigator (PI), Dr. Djorgovski, and his collaborators are conducting a dual effort under the AISRP. They are developing a simple, efficient, user-friendly, interactive package called STATPROG for multivariate statistical analysis of relatively small data sets. They are also developing a large, complex system called FRITZ to help process and analyze large amounts of data (about 3 Terabytes) from the Digitized Second Palomar Sky Survey (POSS-2).

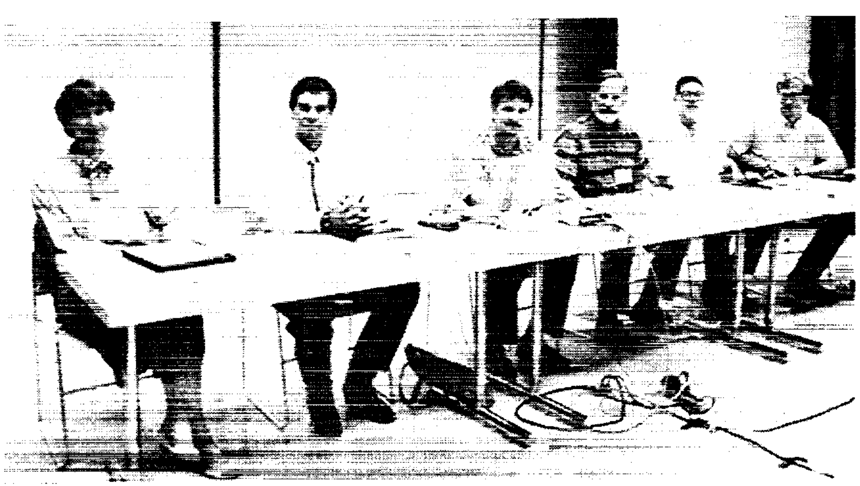
Dr. Djorgovski described the approach taken for STATPROG and its current status. About fifteen stand-alone programs exist and work, and several more are under development or testing. The package has been exported to two external sites for independent testing. Several papers based on the use of the package have been published or are in press. The remaining tasks are mostly documentation in nature. The intent is to deposit the entire package plus the documentation in an anonymous file transfer protocol (ftp) account for pickup by any interested parties.

The FRITZ system is being developed in collaboration with the Jet Propulsion Laboratory (JPL) Artificial Intelligence (AI) group. It now runs on a Sparcstation II under the Sun Unix Operating System, but may be ported to a faster machine shortly. The system will detect objects, measure their properties, classify them, and catalog them. It will use the external charged coupled device

ORIGINAL PAGE
BLACK AND WHITE PHOTOGRAPH

Workstation demonstrations provided an opportunity for interaction between workshop attendees and tool developers

Discussion on key topics and issues continued on Wednesday evening at the NCAR reception



Panelists from the science disciplines discussed science needs and issues with workshop attendees and computer scientists

Assistant could perform. The project was started over a year ago, but most of the work has been done over the past eight months. Removal of cosmic rays from CCD data is a prime challenge, and is the current focus of the Expert Assistant. The initial system has been developed and is being critiqued by the lead users. Based upon their input, the system will be revised and retested with new groups of users. Distribution is planned for the first quarter of FY 1994. There were some questions and discussion regarding implementation of this tool.

SIGMA: Scientists' Intelligent Graphical Modeling Assistant
Dr. Richard Keller
Ames Research Center

This research project is co-funded by OSSA and OAST. Model-building is essential for scientific advancement, yet it is time-consuming and error-prone. The goal of this project is to build a specialized software tool to assist in scientific model-building. The tool has been developed, but has not yet been deployed to support scientists in the Cassini and EOS missions.

The two models being worked with are TGM and Forest-BGC. Dr. Keller discussed why these models were selected, and how the SIGMA approach differs from the typical manual approach to model-building. The three basic differences are that the knowledge base is made available to the system, the scientist conducts an interactive process at the terminal, and the result is an executable computation plan. As an example, Dr. Keller used the Titan atmospheric modeling from the Voyager Flyby. SIGMA developed a model for computing these elements. Details on this process are contained in the presentation charts in Appendix E.

In FY 1993, the project plans to finish the Titan temperature determination, implement the Forest-BGC modeling scenario, build a graphical equation-entering facility, and deliver the first prototype to science users for testing and evaluation.

In discussions, Dr. Keller noted some limits of this tool (the environment doesn't handle partial differential equations). In response to questions, Dr. Keller described the applicability of the tool, the testing by users, and the possibility of trying the tool on other models.

(CCD) imaging for calibration of both the photometry and object classification. Long term plans include the exploration and development of astrophysical research involving very large databases and the use of rapidly developing expert systems. In response to a question, Dr. Djorgovski indicated that there are plans to reanalyze previous sky surveys using these tools.

Multi-Layer Holographic Bifurcative Neural Network for Real-Time Adaptive EOS Data Analysis
Dr. Hua-Kuang Liu
Jet Propulsion Laboratory

Dr. Liu discussed the recent major technical achievements in this investigative area, as well as new technology innovations and publications over the past year or so. His presentation consisted of a theoretical discussion of the basic model and the principle of bifurcation, and the actual results of the experiment. Details of the theoretical discussions and the experiment set-up and results are contained in Appendix E.

This research group has introduced a new class of optical pattern recognition technique which utilizes the bifurcating diffraction phenomenon in non-linear gain saturation memory media. Future research for the bifurcating optical pattern recognizer (BIOPAR) includes the improvement of the quality and resolution of the bifurcating signals and the application of the BIOPAR for Earth Observing System (EOS) data classification problems.

Development of an Expert Data Reduction Assistant
Dr. Glenn Miller
Space Telescope Science Institute

In this project, the investigative team focussed on the data reduction problem, rather than on data analysis or visualization. Dr. Miller described several different types of data reduction and analysis systems and their advantages and disadvantages. The Expert Assistant is an alternative approach which builds on the foundation of these systems and mitigates some of their disadvantages, particularly problems associated with data management and the iterative nature of the work. The goal was to develop a system which will act much as a graduate student: follow instructions on how to do data reduction; check for processing and data quality problems; work longer hours; and have flexibility for modification for new cases.

Dr. Miller discussed the salient feature of the Expert Assistant and how best to involve the scientific community. Their project picked a group of "lead users" who had a real need that the Expert

OAST AI Program
Dr. David Thompson
Ames Research Center

Dr. Thompson (standing in for Dr. Friedland of ARC) gave an overview of the current OAST AI Program. This is a $13M per year program involving seven NASA centers. Earth and space science domains are a major focus of the program at ARC and JPL. The OAST AI Program includes the following projects: Intelligent Scientific Laboratory Instruments; Bayesian Data Analysis--AutoClass; PI-in-a-Box; Multi-Agent Planning for Heterogeneous Registration; Reactive Planning, Scheduling, and Control; Spacecraft Health Automated Reasoning Prototype; and Scientific Analysis Assistant. Dr. Thompson briefly described each of these projects and the current status (See Appendix E for presentation details). All of these projects are intimately tied with users from the very beginning.

DATA COMPRESSION/ARCHIVING/ACCESS/ANALYSIS SESSION

Dr. Randal Davis chaired the session on data compression, archiving, access, and analysis.

Parallel Algorithms for Data Compression
Dr. James A. Storer
Brandeis University

Dr. Storer gave a brief overview of applications of data compression and distinguished between lossless data (decompressed data is identical to the original) and lossy data (decompressed data may be an approximation of the original). He discussed systolic algorithms for lossless compression, lossless compression hardware, image compression visualization tools and experiments, image compression hardware, video displacement estimation, and real-time video compression hardware. However, most of the presentation focussed on vector quantization for image compression and on-line adaptive vector quantization. Dr. Storer showed slides of examples of compression on a medical image at 5:1, a NASA image at 10:1, and a visual portrait at 10:1. The compression did a surprisingly poor job on subtle shading on the visual portrait, and demonstrated the problem with visual quality on this type of image. Details of his presentation are included in Appendix E.

A splinter group chaired by Dr. Storer was set up to further discuss the issues of data compression.

Performance and Scalability of Client-Server Database Architectures

Dr. Alex Delis
University of Maryland

Dr. Delis discussed ADMS, an enhanced client-server database architecture with incremental gateways to heterogeneous relational data base management systems (DBMSs). Today's needs are: inter-database querying, downloading and downsizing, inter-database dependency tracking and change propagation; version and change control; interoperability of heterogeneous relational DBMSs, and multi-site transaction management. There are several technology trends that will affect the development of DBMSs, and a major question is how to take advantage of these technology developments to create DBMSs that offer fast response time and high throughput.

Dr. Delis described the main features of the ADMS system and how it meets the user needs. The prototype platform for ADMS is Unix on Suns, DecStations, and Vaxes, with gateways for Oracle, Sybase, and Ingres. The ADMS enhanced client-server provides distribution of both processing and data, site autonomy (except for updates), and minimal net traffic and overhead. Future work on the ADMS architecture will include: gateway query optimization; pipeline algorithms for interdatabase queries; adaptive update propagation strategies; multi-site transaction management and recovery in autonomous databases; an experiment with increment updates of mirrored databases; and applicability of the same techniques in a multi-processor environment with or without shared memory.

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

Dr. William Emery
Colorado Center for Astrodynamics Research

The goal of this Testbed was to develop on-line AVHRR data distribution including collection archiving and software, emulate some EOSDIS functionality, and prepare snowpack and VI composite images. This is now called the "EOSDIS Testbed" system.

Dr. Emery described the system--how it is accessed, its current features, and plans for the future. The software is available for distribution for Unix and MacIntosh, and PC display software has recently been added. The user base has expanded dramatically. There are about 734 users, and the user base is growing at an average of about four per day. The majority of users are U.S. educational institutions, but there is also a considerable government component. Dr. Emery showed a full resolution data example of what they started with. The image coverage was expanded, which created a lot of network overhead. The Navorder system was created for the users. The usage statistics indicated that users began using the browse feature in a manner that eliminated some of the need for data orders.

The project and users would like to see this system continue until Version 0 of EOSDIS can take over.

Geographic Information System (GIS) for Fusion and Analysis of High-Resolution Remote Sensing and Ground Truth Data
Dr. A. Freeman
Jet Propsulsion Laboratory

Three overflights of the Flevoland calibration/agricultural site were made by the JPL AIRSAR, and the modified VICAR/IBIS GIS was used to analyze the data. VICAR (Video Image Compression and Retrieval) is a set of programs and procedures designed to facilitate the acquisition, processing, and handling of digital image data. VICAR/IBIS is a VICAR-based GIS which requires that all image data be co-registered to a georeference image. The modified VICAR/IBIS GIS is an extension which replaces the "tabular" file format with an "info" file format. The objectives of the modified VICAR/IBIS GIS are: to handle data in many different formats and from many different sources; link all data together through a georeference image; and track data in time, covert pixel values to "actual" values, plot graphs, generate training vectors for classification algorithms, and compare actual and measured parameters.

Dr. Freeman described the implementation and results of the system (details are included in the presentation material in Appendix D). A working GIS/Image Processing System has now been integrated. GIS has been exercised using multi-temporal data from the boreal forest test site and an agricultural site. The rain forest site has been classified using radar data alone. The University of California, Santa Barbara (UCSB) continuous and discrete canopy models have been integrated, and a sensitivity analysis of the models has been conducted. A software tool, "Light-table," has been installed for interacting with very large images. A model has been developed which estimates scattering mechanisms from radar data, and MACsigma0 software has been developed for release to AIRSAR data users. Future activities include entry of ground truth data and correlation with image data; complete integration of the UCSB model with GIS; analysis of boreal forest data; development of a new technique for feature selection/classification, registration of varying terrain height data, and development of a rain forest site vegetation cover map.

Envision: An Analysis and Display System for Large Geophysical Data Sets
Dr. Kenneth Bowman
Texas A&M University

Envision consists of a metadata browser, a data management system, a set of links to feed data to existing visualization tools, and a set of custom designed visualization, analysis, and data manipulation tools. Envision requires regular nD grid type data which must be stored in net common data format (CDF) files. The grids may contain missing data or undefined regions.

Dr. Bowman described the Envision system layout. The project has been concentrating on data files, user interface, and building up the Envision data manager. Envision is used as a metadata browser of netCDP files and as a metadata editor. The user interface and data manager allow changes and edits of metadata. The data manager also manages relationships between files and provides transparent access as a single entity to a dataset consisting of multiple files. Currently, most of the data management features are working, and user interface is being connected. In late 1992, the project expects release of a system with Envision data management facility, Envision interface, and customized connections to NCSA XImage, NCSA Collage, and some NCAR Graphics utilities.

The following presentation on the Spatial Analysis and Modeling System (SAMS) was added to the agenda. This system is a practical application of information systems research.

A Global Satellite Data Acquisition and Analysis System to Support Hydrological Modeling and Regional Climatic Change Impact Studies
Dr. Fran Stetina
NASA/Goddard Space Flight Center

Dr. Stetina's presentation gave on overview of the SAMS for environmental management. Most of the work to date has been in disaster management and planning. Dr. Stetina described the SAMS system and its design: source of data, integration of data, production of products, and distribution of products to users. The entire system is contained in one computer or group of computers networked together. Currently, it is in use in a number of countries, and a demonstration product is being used by the Corps of Engineers using satellite data, along with rain gage data on ground radar, to create more accurate rainfall maps.

Key features of this system are: the use of expert systems to interpret satellite data; a classification model; and an archive manager which allows the model to go in and find the data it needs to run the model. The system is completely automated from receipt of data to generation of models.

Trends in Data Formats for the Space and Earth Sciences
Dr. Randal Davis
Laboratory for Atmospheric and Space Physics

Dr. Davis discussed the changes in archiving, distribution, and use of scientific data which result in increased need for better data formats. In the 1980's, NASA tried to move in this direction with pilot systems such as the Pilot Climate Data System (PCDS), the Pilot Land Data System (PLDS), the Pilot Ocean Data System (PODS), and the Pilot Planetary Data System (PPDS); and operational systems such as the Astrophysics Data System (ADS), the NASA Climate Data System (NCDS), the NASA Ocean Data System (NODS), and the Planetary Data System (PDS). The 1990's will bring the consolidation of NASA's Earth-oriented discipline data systems (NCDS, NODS, and PLDS) into the EOSDIS.

In June of this year, NASA held an invitational workshop to begin to determine if modern data formats will meet the needs of the future. A number of different formats were discussed at this workshop, and it was concluded that good data formats are available for space and Earth science applications, but the relationship between special scientific data formats and data formats from general computing has to be examined further. As a result, developers of CDF, netCDF, and hierarchical data format (HDF) are examining the possibility of developing a common interface to data in their formats. Dr. Davis indicated that the next workshop will be sometime in the Fall.

There was some discussion among the AISRP workshop attendees on the formats and differences within the community, and how users interact with this issue. A splinter group session was set-up to further discuss this topic and develop some recommendations.

Wednesday, August 12

SCIENTIFIC VISUALIZATION SESSION

Dr. Mike Botts chaired the session on scientific visualization. Most of the presentations in this session were accompanied by visual demonstrations via computer projection screen.

The Grid Analysis and Display System (GrADS)

Dr. James L. Kinter
Center for Ocean-Land-Atmosphere (COLA) Interactions
University of Maryland

Dr. Kinter provided background information on the COLA group at the University of Maryland, and the types of research in which they are currently engaged. Based upon a survey of users conducted by Dr. Mike Botts, a number of complaints about existing scientific visualization tools were summarized. Most users are not interested in the software tools currently available for these reasons, which were the same problems identified by the COLA group. As a result, the COLA group decided to develop GrADS. They have tried to tightly integrate the following: access of the data; manipulation of the data interactively; and display of the data to the scientist in a way that is familiar. Other goals of this system were interactivity, ease of use, and the capability to generate a hardcopy output.

Dr. Kinter described how GrADS was designed to meet these goals, how it addressed the disadvantages noted in the survey, and the hardware requirements for the system. An on-line demonstration was conducted, and an example of hardcopy output was shown. This system has been introduced to a group of users, and considerable feedback has been received which will enable future improvements.

National Center for Atmospheric Research (NCAR) Interactive Status

Dr. Bob Lackman
NCAR

Dr. Lackman discussed the collaborations with other groups since last year's workshop. They are using GrADS at NCAR, and with NCAR graphics, are concentrating on supporting the community NCAR graphics package with documentation and user support. The first NCAR graphics user conference, June 17-19, included hands-on training using workstations as well as presentations, demonstrations, and panel discussions. Dr. Lackman provided on overview of conference topics, and made hardcopies of the presentations available to workshop attendees.

The NCAR interactive community goals are: to guarantee long-term support via cost recovery; provide university and non-profit researchers low cost visualization; and advance and support the scientific infrastructure through common software. NCAR graphics builds on existing libraries, is a single package for distribution, and is portable across systems. The package is set up to have three levels of interface: programmatic; command line; and visual point and click. Dr. Lackman described what NCAR interactive will look like. Currently, they have a functional requirements document, have completed preliminary design, and have a prototype high level utility. Next year, they will have a command interpreter, and the year after that, they will be distributing software as part of the NCAR graphics distribution.

<u>A Distributed System for the Visualization and Analysis of Observed and Modeled Meteorological Data</u>
Dr. Steven Koch
Goddard Space Flight Center

The goal of this project, a joint effort of GSFC and the National Center for Supercomputing Applications (NCSA) was to create a tool for handling the large amounts of data generated by satellites, observational field programs, and model simulations; and to extend existing 2D mapping capabilities with new analysis functions and modern techniques of 3D visualization, user interaction, and animation. Dr. Koch discussed the project approach and its results. The system, called GEMVIS, maximizes the use of existing software; uses commercially available visualization and application builder tools; provides visualization and analysis capabilities in the areas of 3D volumes of data, evolution of data over time, and distributed processing; and provides a highly interactive environment on a single display.

Dr. Koch and Dr. John Hagedorn conducted a demonstration of the GEMVIS using the Explorer map module and functions, deriving vorticity display. The demo included other orientations and 3D renderings, as well as animations of wind vectors. They discussed the accomplishments at GSFC and NCSA, as well as problems encountered and the gains and losses of 2D GEMPAK. Near term future work includes minor user interface improvements and fixes, preliminary annotation, image loop animation, and user documentation. They will release the software for use in the Sever Storms Branch of GSFC in October 1992. Long term future work includes a database for metadata, user interface enhancements, additional visualization techniques, and additional animation.

There were questions and discussions related to the timing of distribution of modules and portability to systems other than SGI. Currently, they are now distributing it on Sun and Cray, and are looking at a six month window for distribution on HP and DEC. The problems with proprietary systems was noted and discussed.

<u>Planetary Data Analysis and Display System: A Version of PC-McIDAS</u>
Dr. Sanjay S. Limaye
University of Wisconsin-Madison

McIDAS is an evolutionary hardware/software system for Earth Atmospheric data in use since the mid-1970's. The planetary version is aimed for analysis of primarily imaging data from space missions such as Voyager, Pioneer Venus, Magellan, Hubble, Mars Observer, and Cassini. McIDAS-X is a version for Unix workstations (RISC-6000, SGI, Sun) with X-Windows. Planetary PC-McIDAS implementation was begun under McIDAS-X. Dr. Limaye described the current status of system level development, new software applications, the data descriptor block, multiband data display and processing, and multispectral analysis. Dr. Limaye discussed and demonstrated how the optical navigation technique was used to successfully remove an error in roll angle to correct Voyager Neptune images.

<u>Experimenter's Laboratory for Visualized Interactive Science</u>
Dr. Elaine Hansen
Colorado Space Grant Consortium

This project was a group effort among the Colorado Space Grant Consortium, LASP, NCAR, GSFC, and UCSB. Dr. Hansen presented an overview of the program and its interactions with the user community. The program goals are to provide a capability that helps visualize data to better understand the large, complex, diverse, and multi-dimensional data sets; support science research within and across NASA science disciplines; to provide a laboratory that can be easily used and tailored; to provide tools at an affordable price; and to capitalize on existing systems techniques, technologies, and tools.

Dr. Hansen introduced Dr. Margi Klemp, who discussed the usability analysis and software design. Usability challenges include: the integration of visualization, data processing, exploratory analysis, and communication needs into a single application; a user interface which will hide the complexities of the hardware and software required to enable these functions; the flexibility to make changes dynamically in the applications and to add user specific functionality; input of diverse data formats; annotation; and direct manipulation interfaces for interaction with data. Dr. Klemp reviewed the project accomplishments and status to date. Future work includes completion of the Alpha version by the end of the year, the continuation of user testing, the integration of the spreadsheet engine, the development of direct manipulation user interface, new graphics capabilities, enhancement of TAE+, hooks for calling external functions, development of a direct manipulation annotation editor, color hard copy interface, and the creation of an interface to GL for hardware rendering.

Dr. Bill Boyd showed a slide demonstration of the PolyPaint functionality, which renders datasets as 3D surfaces. This system provides the capability to look at several features at once, and can facilitate meteorological research. The project is trying to get these tools into scientists' hands to enable them to incorporate this into the research cycle.

<u>SAVS: A Space Data Analysis and Visualization System</u>

Dr. E. Szuszczewicz
Laboratory for Atmospheric and Space Science
Science Applications International Corporation

SAVS is a combined effort among SAIC, Advanced Visual Systems (AVS), and the University of Maryland. The focus of this system is on the multi-disciplinary databases designed to understand the cause-effect relationships in the solar-terrestrial system and their extrapolations to other planetary bodies. The major components of the system are: innovative visualization software (AVS); advanced database techniques; a set of mathematical, analytical, and image processing tools; and a strongly developed sense of scientific requirements.

Dr. Szuszczewicz discussed the needs of the practicing scientist and how the SAVS attempts to meet those needs. The visualization system is wrapped around AVS, which provides a variety of tools for rendering volume data. However, AVS is just one component in the overall system. The SAVS design goals focus on ease and functionality. Dr. Szuszczewicz discussed the interactive functionality of the SAVS system and demonstrated this capability on-line. During this first year of funding, the AVS has been ported to lower-end platforms, and an extensible user-friendly architecture and data and model interface modules have been developed. Basic mathematical and statistical functions have been implemented, and the development of hooks for an interactive interpreter has been started. The system has been tested on the Combined Radiation Release Experiment Satellite (CRRES) and International Sun Earth Explorer (ISEE) orbits and local data bases, and plans have been initiated for remote data access capabilities. In response to a question from the audience regarding documentation, Dr. Szuszczewicz indicated that there will be a year-end report as will as NASA press releases and science applications reports. There was some discussion relative to strategies in working with vendors to accommodate proprietary system issues. Mr. Bredekamp noted that one of the goals of the program is to engage the vendor community and work licensing agreements for the benefit of science users. Some of the program participants are currently working with vendors on this issue.

NASA Supported Advanced Visualization Techniques

Dr. Bill Hibbard
University of Wisconsin-Madison

Dr. Hibbard discussed VIS-5D and VIS-AD. VIS-5D (VISualization of 5-Dimensional data sets) is used to visualize large output data sets from numerical weather models. This tool is about four years old, and is currently used by scientists at the University of Wisconsin (UW), NASA/MSFC, NASA/GSFC, and other sites. It runs on SGI, IBM RISC, and Stardent, and is available as freeware by anonymous ftp. VIS-AD (VISualization for Algorithm Development) is a visual laboratory for experimenting with algorithms for extracting useful information from remote sensing data. It puts visualization where people work on developing algorithms for automatic processing.

Dr. Hibbard discussed how the user can utilize VIS-AD in a highly interactive environment. He showed how VIS-AD is used in construction of algorithm. As part of his presentation, he gave a demonstration of VIS-5D and its interactive capabilities. VIS-AD can also be viewed as a McIDAS macro language where the user can invent data structures as needed, display any data object easily, and invent display frames of reference. VIS-AD can access McIDAS data structures. Dr. Hibbard noted that anyone who is interested in using this tool can contact him for instructions on how to obtain it.

LinkWinds: The Linked Windows Interactive Data System

Dr. Allan S. Jacobson
Jet Propulsion Laboratory

The objectives of this project were to: develop a software environment to support the rapid prototyping and execution of data analysis/visualization applications; provide a suite of tools to interactively visualize, explore, and analyze large multivariate and multidisciplinary data sets; develop a user interface which allows maximum data and tools accessibility with a minimum of training; and provide system design and tools to make the environment accessible to application development by users. Dr. Jacobson described the system and how it is implemented, emphasizing that the scientists drive the development of the system.

LinkWinds has standardized on HDT, but other data formats are accessible via interaction with DataHub. A user has to edit only three files to use the database. Dr. Jacobson conducted a demonstration of LinkWinds from ozone and water vapor data from the Microwave Limb Sounder daily files. Currently, LinkWinds has only been distributed to a few locations outside of JPL. More sophisticated tools are needed and are planned. Future plans are to port LinkWinds to other Unix platforms, expand the standard input data formats via interaction with DataHub, and develop applications for visual data selection and subsetting. They also plan to implement an applications generator to support user applications development, and affect a wider distribution of LinkWinds.

DataHub: Knowledge-Based Science Data Management
Dr. Tom Handley
Jet Propulsion Laboratory

Dr. Handley discussed the key data base issues which had been identified by Dr. Jeff Dozier of the EOS Program. Scientists want to query and analyze a variety of different types of data. Analysts and modelers need access to stored satellite data, derived products, and model output on numerous and large objects, with data distribution among nine archive centers. Researchers want to render data on local workstations to conduct their research. There are a number of challenges to meeting all of these needs. DataHub is a value-added, knowledge-based server between the data suppliers and the data consumers. It meets the need to address the barriers associated with distributed, autonomous, heterogeneous systems.

Dr. Handley described the relationship between DataHub and LinkWinds, and discussed the functional architecture, the data model, the software architecture and implementation, and the user created datasets. To date, the project has defined a general framework for science data management, identified a critical subset of data operations for the science data visualization applications, and developed an initial prototype (DataHub 0.3) with common user interaction, data format conversions, user inventory management, a DataHub/LinkWinds interface, and underlying object-oriented structure and methods. After Dr. Handley discussed current status and future plans, there were some questions regarding how this could be used in the EOS system, and some interest was expressed in bringing this before the EOS advisory committee.

Visualization Discussion
Dr. Michael Botts
University of Alabama in Huntsville

Dr. Botts has been on a special assignment to NASA Headquarters as a result of Congressional attention regarding visualization on the EOS mission. He made his final report, "The State of Scientific Visualization with Regard to the NASA EOS Mission to Planet Earth," available to the meeting participants, and presented its findings as well as some options for meeting the visualization requirements of the EOS mission.

A survey of Earth scientists was conducted relative to problems with current visualization tools. The applications of visualization are: scientific investigation; data validation; model and algorithm development and validation; data browse; information transfer; and mission operations. The primary use for visualization is for scientific investigation. Dr. Boots discussed the variety of reasons why scientists are not using the capabilities that are available to them today. Most of the visualization bottleneck is related to lack of adequate software which allows the scientist to take advantage of hardware power and to interactively visualize and analyze the data. It is questionable whether Commercial-Off-The-Shelf (COTS) software will be adequate for meeting all the needs of EOS. Dr. Botts discussed the advantages and disadvantages of COTS, public domain, and in-house development software.

With regard to visualization, the development environment within OSSA can be characterized as fragmented, and lacking adequate organization and funding structures as well as adequate mechanisms for technology transfer both within OSSA and between OSSA and OAST. Dr. Botts identified several general areas presently needing additional consideration and concentration. His report concluded that visualization is vital for meeting the scientific objectives of the EOS mission, and the applications software for putting visualization techniques and capabilities into the hands of the scientist are at present inadequate. Although there is increasing probability that COTS software can serve as a core for meeting many visualization needs, there will still be a need for in-house development efforts focussed on extending or modifying these tools to meet application-specific requirements.

Dr. Botts presented some ideas on what could be done to increase the effective use of visualization and analysis tools. Items discussed were: the possibility of a visualization/analysis working group; assistance centers, located at one or more Distributed Active Archive Centers (DAACs); use of pathfinders; vendor programs to improve the success of COTS; improvements to the licensing/procurement process; and improvements to publishing and remote interaction.

In the discussion period, it was noted that all of the comments are valid for other disciplines as well as the Earth science and applications field. Inasmuch as there is natural resistance to doing the "learning" required to utilize some of these tools, it would be worthwhile to have support for the learning cycle. An infrastructure like this workshop could facilitate discussions on how to solve problems. There was some discussion on the "assistance centers," which were generally seen as a good idea if they could be staffed with diverse expertise. It was felt that collocating an assistance center with a DAAC would be essential to making it work. One of the challenges will be forging a strong link between the Information System Branch program and the EOS program. Other aspects discussed by the workshop participants were: the "right way" to do software support; commercialization, the technology transfer issue and incentives to creative development; the problem of getting scientists to use the tools; mechanism for getting support for multi-discipline, non-project specific aspects of the program.

One of the suggestions made for getting tools out to the various communities was for the participants in this group to go to other external well-established society meetings, such as the American Geophysical Union (AGU) and give presentations and demonstrations. Publication in professional journals was another suggestion. Another possibility that was discussed was a NASA Research Announcement (NRA) for investigators to propose applications for existing tools. These would provide some "seed" money to try out the tools within the scientific community. This type of mechanism is already a part of the EOS program. Another idea was for a "consumer's guide" on current tools, directed to the scientist.

Thursday, August 13, 1992

Before beginning the morning session, the Chairman Mr. Mucklow announced adjustments to the agenda for the day.

PROGRAMMATIC ISSUES AND ANSWERS SESSION

<u>Information Systems Research and Technology Reports</u>
Mr. Glenn Mucklow
NASA Headquarters, OSSA

The objectives of this program are to apply advanced information systems technology to improve support to OSSA science programs, and enable continual evolution of OSSA data systems environment and supporting infrastructure. The elements of this program include the Investigator Working Groups (IWG), applied research, technology development, and systems evolution. Mr. Mucklow discussed the approach for each of these elements. The technology transfer element is the key to the program. It includes testbeds, software support, strategic users, commercial aspects, and COSMIC, and seeks to bridge the gap between research and technology development and mission operations and post-mission science research.

Electronic Mail and News Groups
Mr. Randy Barth
Goddard Space Flight Center

Mr. Barth provided an overview of the NASA POBox and how to use it. The NASA POBox simplifies mail exchange among AISRP researchers on various systems (the Internet, NSI-DECnet, RSCS Networks, X.25 Networks/Nodes, and LAN & PCs connected to these systems). Currently, there are three AISRP group distribution lists: aisrp-members, for all members; aispr-pi for all PIs; and aisrp-general, for general discussions. The NASA Science Internet (NSI) Help Desk can answer questions on the use of the NASA POBox.

There are several mechanisms for electronic discussion groups: discussion groups vis a central bulletin board; discussion groups via USENet NEWS, and discussion groups via Electronic Mail. Mr. Barth discussed the advantages and disadvantages of each of these mechanisms.

NASA Science Internet Developments
Dr. Christine Falsetti
Ames Research Center

Internet, the largest Federal network, is a collection of various networks. The NSI was established to provide communications to NASA's OSSA. It provides computer networking services, management and operations support, and technical assistance to authorized users throughout NASA centers and research institutions worldwide. NSI's goal is to provide a high-speed communications network that connects all space scientists, providing ready access to data and information stored anywhere in the world.

Dr. Falsetti described the current telecommunications infrastructure and some of the evolving network-based applications. The plan for evolution is to use the networks in place and use advance technologies to go from 45 mbps to 2488 mbps. The National Research and Education Network (managed out of OAST) is leading this activity. NASA and the Department of Energy (DOE) have collaborated in a procurement to obtain advanced technology, which will enable universities to have better access to the service. In response to questions regarding connections to service, Dr. Falsetti advised participants to contact their regional service and find out what their plans are to connect to the network. NSI will serve as the backbone to EOSDIS.

There was some discussion regarding the funding of regionals, and the prospective commercialization of these services, and how university users would be affected. Dr. Falsetti indicated that there are potential problems in this area, and this issue is currently in heavy debate. Further discussion on this subject was deferred to a later group discussion.

CESDIS: Center of Excellence in Space Data and Information Sciences

Dr. Ray Miller
NASA/GSFC

The CESDIS mission is to bring together computer scientists from university, industrial, and government laboratories to: conduct computer science research having application to Earth and space science; focus attention on accessing, processing, and analyzing data from space observing system; and collaborate with NASA space and Earth scientists. CESDIS is managed by the Universities Space Research Association (USRA). Dr. Miller described the CESDIS organization, activities and tasks, and the research currently being conducted by the CESDIS staff.

The Stanford University project has been developing software to automate the analysis of global auroral images obtained from Dynamics Explorer (DE)-1 and Viking satellites. The Duke University project is involved with parallel compression of space and Earth data. The AISRP Projects are being reported upon at this Workshop. Dr. Miller described other additional research tasks, as well as consultants and fellowships that are being funded through CESDIS. CESDIS has also conducted annual workshops on various data and information systems topics. CESDIS provides support to peer reviews for NRA's, organizes the "Advances in Computational Sciences Seminar Series," and provides support to the Minority University Space Interdisciplinary Network project. Future CESDIS direction include more coordination of AISRP projects, assistance with Earth and space science HPCC projects, and the building of ties with EOSDIS.

SCIENCE DISCIPLINE DISCUSSIONS

Dr. Amy Walton chaired the session on the science discipline discussions. Panel members were: Dr. Tom Ayres, Astrophysics; Dr. Mike Botts, Earth Science; Dr. Jim Willett, Space Physics; Dr. Robert Jackson, Life Sciences; and Dr. Steven Lee, Solar System Exploration. Each panel member made brief opening remarks prior to the general discussions.

Astrophysics

Dr. Tom Ayres

Astrophysics has an unusual type of data set--the images are relatively sparse and easily compressed. Except for solar physics, the data sets are not large. However, there are many different platforms and missions, and many different interfaces into the data set as well as a ground-based connection. In terms of bottlenecks, one of the major concerns is that NASA is not providing enough funds to procure new workstations, which are needed for more disk space and faster CPU. One of the most valuable tools is a screen editor, and the next most valuable is a flexible software environment. To implement research, scientists in this discipline need transparent access to existing and new data sets, access to archival data, and easily modifiable software. In parallel, they need access to expertise and specific software (or software modules). The researchers do not need a proliferation of sophisticated visualization systems, but tools that can be easily modified to attack specific problems.

In the panel discussion, Dr. Ayres emphasized that the need for interdisciplinary science sharing has increased, and the PDS has had an influence on astrophysics. In astrophysics, the primary science problem is not enough telescopes in orbit. The problems some disciplines are encountering with large data sets has not been an issue for astrophysics.

Earth Science
Dr. Mike Botts

The major issues within Earth science involve large data sets, and lots of data, and interuse of multiple data sets from multiple disciplines. In large data sets, the issues are: keeping interactivity in visualization (adaptive sampling could help); automatic feature recognition and tracking; and data compression (lossy vs. lossless - how much is acceptable). With the restructuring on EOS, there is less capability to do on-board processing. In the interuse of data sets from multiple disciplines, there are navigation, gridding, and projection issues. Dr. Botts recommended leaving gridding and navigation till as late as possible, but tools will be needed to do this. The advantages and disadvantages of adaptive sampling were also discussed.

The general discussion centered around data formats. A suggestion was made to have a limited library of formats for choice by users, and visualization tools would be responsive to this limited library of formats. It was noted that equal emphasis is needed on the analysis process, and one approach is to get scientists involved very early in the process. Dr. Mucklow indicated that there is a mechanism for doing this.

Space Physics
Dr. Jim Willett

Dr. Willett indicated that this workshop had been very useful and helpful, particularly the opportunities to talk with other participants and share ideas. He noted that the tools produced by the scientist with the help of computer scientists are the ones that are most successful and are actually being used in research. It would be useful for someone to put together a "consumer report" document to categorize all of the tools/systems, describe salient features, and show how they fit together. The workshop participants need such a document to go into their discipline communities and talk about what's available.

With regard to the earlier discussion on the problem of getting tools into the community, Dr. Willett noted that in space physics, the younger investigators used and recognized the value of new tools, particularly visualization tools. Some of the tools demonstrated at this workshop might be very useful in displaying magnetospheric modeling. One possible approach might be to initiate a joint-funded activity (space physics and information systems) to pursue implementation of these tools, which could also help with technology transfer. A suggestion was made to consider having a technology transfer supplement to the education outreach grants.

In response to a question, Dr. Willett stated that the National Space Science Data Center (NSSDC) is a critical element in the space physics and other discipline data systems. It would be a node on the space physics data system, and provide a master catalog and a deep archiving capability. There was some discussion of the current OSSA data policy, and the evolution to this policy by all space science disciplines. Currently, space physics is the discipline most firmly entrenched in the old system (data proprietary to the PI for a period of time). It was noted that there is current disarray on how this issue is handled in OSSA.

Life Sciences
Dr. Robert Jackson

Dr. Jackson described the Life Sciences program for Space Station Freedom, and two basic research scenarios--automatic operation and crew assisted operation. The automatic operation mode provides routine operation of major facility equipment to support biospecimen growth and development. The bottlenecks in this mode are primarily limited downlink bandwidth, downlink interruptions, and no on-board communications outage recorder. Researchers need tailored bandwidth reduction techniques and efficient on-board data processors and storage. The crew operated mode provides for crew to set up runs, manipulate specimens, collect data, collect and analyze samples, and maintain equipment. The bottlenecks in this mode include crew time and skill maintenance, limited uplink and downlink bandwidths and interruptions, late delivery of data and samples, and long storage time of samples. Research is needed to enable effective conduct of advanced experiments to accommodate limited uplink and downlink, minimize the demand for power and volume, and maintain and improve crew skills.

In response to a question about metadata, Dr. Jackson that it has been inadequate, and life sciences is just starting to begin an archive. They will need help with this development, and the implementation to bring it on-line. Recoverable archive remains one of the major problems. Tools are needed to store more information up-front. The PI-in-a-box concept was discussed as a potentially useful tool for life sciences.

Solar System Exploration
Dr. Steven Lee

Dr. Lee provided a brief background of the PDS and its structure. JPL provides the management and top-level catalog. Colorado is in the Atmospheres Node. Each discipline node divides up among specialty nodes. Problem areas in archiving planetary data from active missions are: the obligation of flight projects to archive data; preservation of mission funds for data archiving; availability of project personnel for archiving tasks; definition of interfaces between projects and PDS; large lead times needed to influence archiving plans; selection of storage technology; massive data volumes; and the proprietary period and scheduling of data transfer. PDS is actively working all of these areas.

The data formats issue came up again during discussion of the PDS. Dr. Davis added that the working group has come to the conclusion that formats should not be prescribed; a more workable approach is to described a standard set of data objects, and encourage users to accept those standards. It was noted that the tools demonstrated at the workshop have applicability to the planetary research discipline. Another issue discussed was how to get rid of old data where the cost of maintaining the data for exceeds its usefulness. Space physics is currently looking into this problem.

TECHNOLOGY TRANSFER DISCUSSION

Dr. Tom Handley and Dr. Larry Preheim led the group discussion on technology transfer.

For the purposes of this workshop discussion, technology transfer was defined as the transfer of organized knowledge to a project or program for the eventual purpose of producing new or improved products, processes, or services. The transfer can occur through consulting, documentation, training, demonstration, or collaborative technical work. Too often research and development has been content to "throw its product over the wall and hope someone will catch it." Technology issues are viewed differently by advanced development groups and implementation/production groups. Dr. Handley discussed some of the barriers to technology transfer. The user community lacks a process to identify common requirements, and a lacks a vehicle to exert leverage on NASA to implement common designs. The resources invested in existing system and applications, as well as the attitude and culture of the work force, make it difficult to evolve to new technologies. In addition, there are inadequate incentives to foster the insertion of new technology into new missions, and there is a fear of not being able to meet mission objectives (performance and budget) using "newer" technology. Overall, there is no documented coherent NASA vision for broad-based technology integration or a technology transfer process.

It was noted that OSSA and OAST have been trying to work out a process to address some of these barriers and develop a coherent NASA vision. One example of good collaborative effort is HPCC; the key was participation and drive from the onset of the project. Dissemination of information alone does not produce results. Technology transfer occurs within the context of one-on-one relationships of technologists and organizations. The programmatic challenge is to establish these relationships. JPL is putting together a process for technology transfer at that institution. It includes a technology transfer readiness review of on-going and new programs. However, the dominant project individuals need to be advocates of technology transfer and keep the user community involved in the process. A flow diagram of the process is included in the presentation package in Appendix D. The advantages of collocation were discussed, and it was generally agreed that directly collaboration, particularly at the beginning of the process, is optimum. An intermediate "technology transfer" group can be used to sustain the activity once it is well started.

Another issues that was discussed at some length was how to support the transfer once it is successful. Support can be time consuming and expensive, and things that need a lot of "handholding" tend to stay in research labs for this reason. Dr. Handley noted that there are several types of transfer: a single, stand-alone tool for a group of users; a tool which is an integral part of a larger system (e.g. shuttle scheduling) to other users; and technology which has only "potential" users.

How to move commercial products into the NASA community was another popular topic, and the participants discussed the issues with licensing. Dr. Botts indicated that he has been working on this issue, and invited interested parties to get in touch with him to discuss it further.

Dr. Preheim presented a view of an "optimal solution:" establish three demonstration sites, GSFC, JPL, and LASP, where prospective users/interested parties could visit, view the tools, and have expertise there to walk them through the products. However, there would need to be coordinated testing for maturity and useability of the technology. One center (JPL) would be the focal point to address commercialization issues. This idea was discussed among workshop participants, who saw value in this type of approach. However, it was noted that users would need to know the availability of products at the center, and a "consumer's guide" was again suggested as a way to fulfill this need.

Mr. Mucklow indicated that by this time next year, the AISRP investigators will need to identify how they will do technology transfer. A suggestion was made to have an archive site (e.g., LASP) for the software that was described and/or demonstrated at this workshop. In response to a suggestion regarding demonstrations in a special session at AGU, Mr. Mucklow asked that anyone interested in doing this, please contact him and it would be pursued.

COLLABORATION AND COMMUNICATION DISCUSSIONS

Mr. Mucklow, Dr. Willett, Dr. Gordon Johnston, and Dr. Peter Friedland led the group discussions on collaboration and communications.

Dr. Johnston, the thrust manager for science in OAST, explained how OAST has been restructured from discipline-oriented management to the focused thrust structure consisting of science, exploration, and operations. Dr. Peter Friedland from ARC talked about the OAST culture, which is significantly different from OSSA--it is one of research of technology to facilitate NASA missions. In OAST it is unusual to have large scale peer reviews--proposals are often unsolicited, and are funded through the field centers. The AI program was advocated and approved as a mission-oriented thrust. The goal in the OAST program is to do a great deal of technology transfer at all levels, and the program office supports collaborative efforts with OSSA--in advocating, funding, and working. With respect to the issue of technology transfer on flight project, Dr. Friedland acknowledged that technology transfer is more difficult. However, technology transfer is easier to accomplish when connected with an experiment instead of part of the flight transportation system (e.g. the shuttle). A good example of this was the PI-in-a-Box.

The group discussed the possibility of more joint NRA's. There was general consensus that this approach would be very attractive to the scientific community. OSSA hopes to increase the CESDIS activity to provide a mechanism for OAST and space scientists to work together. In response to a comment regarding funding for support to space science, Dr. Willett noted that the issue of science support funding has come up repeatedly in OSSA and in the Space Physics Division, but most of the activity in the office is toward getting new projects to fly. Mission operations and data analysis (MO&DA) funding is dependent on flight missions. The office would like to get some sort of statement regarding the right balance among funding to the science community, funding to hardware, and funding to extend missions. Funding was clearly recognized as a key issue, but it was also recognized as something that, to a large extent, is beyond NASA's control, and is driven by national priorities. What the community can do is work through the established advisory committees and have enough small scale missions to provide some support for R&D. The National Academy of Sciences (NAS) and the National Research Council (NRC) are the recognized advocates for science, and the community needs to exert influence through this structure.

Dr. Willett noted that a mechanism for technology transfer of visualization tools could be an NRA type of activity where funding is provided to selected groups to try out the tools and bring in the technology. OAST is trying to reduce the cost of mission operations, which could make some more funding available for science analysis.

OPPORTUNITIES FOR FUTURE AISRP RESEARCH DISCUSSION

Mr. Mucklow led the wrap-up discussion for the workshop. He noted that HPCC is currently under evaluation, and four to five major awards will result. This will generate some massively parallel processing testbeds. In addition, OAST is sponsoring an HPCC CESDIS announcement. OSSA has funding in FY 1994 to fund additional AISRP grants, and will have an announcement coming out in 1993. As a result of these workshop discussion and lessons learned, the program office will try to have more focussed topics of research. OAST could help provide an opportunity for some of the creative proposals that fall outside the scope of the next announcement.

Some key discussion areas identified during the workshop were issues associated with data compression and analysis, data formats, and technology transfer.

Dr. Davis, Dr. Storer, and Dr. Falsetti summarized their splinter group findings on data formats, data compression, and networking.

REPORT OF SPLINTER SESSION ON DATA FORMATS
Dr. Randal Davis

No one data format can do it all, even within a single discipline. Factors to consider in selecting data formats are: usability, performance, ease/difficulty of change, completeness of metadata, and support for data compression. Software must be provided for every data format. At a minimum, it must display any header or label information, create a programming language data structure, and put the data into a "plain" file. A "consumer reports" kind of analysis is needed to evaluate and compare formats. Data formats should be "bottom-up" development. Some concern was raised about "top down" data formats standards efforts, such as the SFDU concept.

Specific needs differ according to the type of data being handled: archive data, active data, quick look/browse data, and operational data. Archive data needs lots of metadata, minimal system dependencies, and lossless compression. Metadata for active data is dictated by need and system dependencies may be acceptable. Quick look/browse may not need much metadata, and system dependencies and lossy compression may be acceptable. Operational messages require minimal metadata and are heavily encoded.

REPORT OF SPLINTER SESSION ON DATA COMPRESSION
Dr. James Storer

NASA needs for data compression are: archiving of data; transmitting data more quickly between NASA centers and research institutions; data distribution on CD-ROM, etc.; fast data browsing. It is clear that with over a tera-byte of data per day to be archived in the future, fast ways to "finger through" the data are critical.

The tradeoff between storing less data or having more data with less fidelity for specific key NASA applications needs to be studied. Better measures of distortion must be developed that can evaluate compression algorithms for specific applications and which allow the user to quantitatively specify a tradeoff between compression and fidelity. Note that such measures are also critical for compression algorithms to adaptively "learn" about the data. Continued research is needed on general purpose adaptive techniques that work well for the wide variety of NASA data and have an easily "adjustable" compression-fidelity tradeoff. Fast algorithms (and hardware) for real-time data compression are needed so that compression and decompression can be "invisible" to the user. Integration of data compression software and hardware into current and proposed NASA archives and distribution systems needs to be studied. Better communication between Earth scientists and computer scientists is needed so that compression systems that are a good "match" to the applications can be developed, and so that Earth scientists are comfortable using these systems.

Gathering data and transmitting data from remote sensors is an important issue for NASA, and one for which compression can play an important role by effectively increasing the bandwidth of communication channels. However, due to the many highly technical issues involved, it is reasonable to view it as a separate problem from the compression of data that has already been obtained.

One action that was suggested by the participants was to identify some data sets on which to do a data compression exercise. This could be worked on at the Data Compression Workshop in the spring. Mr. Mucklow took the action to follow up on this.

REPORT OF SPLINTER SESSION ON NETWORKING
Dr. Christine Falsetti

Issues identified and discussed were:

- Evolving NSFNET infrastructure and support for NASA PIT's at NSFNET regional institutions. What is the impact on NASA PI's? What reponsibility will NASA take to support NASA PI's? What are our plans for support? How does NSI track NASA PI's?

- WAN-LAN roles and responsibilities and interface issues at sites. As WAN upgrades to high performance network delivery, what are the implications at the science site? Communities must be alerted that they should be working with LAN providers to ensure that they will be able to take advantage of the high-performance network when deployed.

- Request for more information about NSI at SWG's, particularly EOSDIS Data Panel meetings.

- How to better listen to the science community. Perhaps volunteers for involvement in restructured scientist input (SSC) activities?

A final action on the workshop participants in general was to pass along any items of group interest via electronic mail; if paper, pass along to Mr. Mucklow for distribution.

APPENDIX B:

Participants and Attendees

APPLIED INFORMATION SYSTEMS RESEARCH PROGRAM (AISRP) WORKSHOP II

Laboratory for Atmospheric and Space Physics
Space Technology Research Building
University of Colorado
Research Park
Boulder, CO

AUGUST 11-13, 1992

AGENDA

__Tuesday, August 11__

7:30 am	Coffee and Registration
8:30	Opening Remarks: Scope of Workshop **J. Bredekamp**/NASA Office of Space Science and Applications **G. Mucklow**/NASA Office of Space Science and Applications

ARTIFICIAL INTELLIGENCE/EXPERT SYSTEMS/NEURAL NETWORKS SESSION
Chaired by **P. Friedland**

9:00	Multivariate Statistical Analysis Software Technologies for Astrophysical Research Involving Large Data Bases **G. Djorgovski**/California Institute of Technology
9:30	Multi-Layer Holographic Bifurcative Neural Network Systems for Real-Time Adaptive EOS Data Analysis **H. Liu**/Jet Propulsion Laboratory
10:00	BREAK and DEMONSTRATIONS
10:30	Development of an Expert Data Reduction Assistant **G. Miller**/Space Telescope Science Institute
11:00	Construction of an Advanced Software Tool for Planetary Atmospheric Modeling **R. Keller/Ames** Research Center
11:30	Office of Aeronautics and Space Technology (OAST) Artificial Intelligence in Earth and Space Science **D. Thompson/Ames** Research Center
12:00	LUNCH and DEMONSTRATIONS

DATA COMPRESSION/ARCHIVING/ACCESS/ANALYSIS SESSION
Chaired by R. Davis

1:30 pm High Performance Compression of Science Data
J. Storer/Brandeis University

2:00 VIEWCACHE: An Incremental Pointer-based Access Method for Autonomous Interoperable Databases
A. Delis/University of Maryland

2:30 A Land-Surface Testbed for EOSDIS
B. Emery/University of Colorado

3:00 BREAK and DEMONSTRATIONS

3:30 Geographic Information System for Fusion and Analysis of High Resolution Remote Sensing and Ground Truth Data
L. Norikane/Jet Propulsion Laboratory

4:00 An Interactive Environment for the Analysis of Large Earth Observation and Model Data Sets
K. Bowman/Texas A & M University

4:30 NASA Data Formats (HDF, CDF, Net CDF, etc.)
R. Davis/University of Colorado

5:00 Adjourn to Demonstrations

Wednesday, August 12

8:00 a.m. Coffee

SCIENTIFIC VISUALIZATION SESSION
Chaired by M. Botts

8:30 The Grid Analysis and Display System (GRADS): A Practical Tool for Earth Science Visualization
J. Kinter/University of Maryland

9:00 An Interactive Interface for NCAR Graphics
R. Lackman/National Center for Atmospheric Research

9:30 A Distributed Analysis and Visualization Systems for Model and Observational Data
M. Arrott/National Center for Supercomputing Applications
S. Koch/Goddard Space Flight Center

10:00	BREAK and DEMONSTRATIONS
10:30	Planetary Data Analysis and Display System: A Version of PC-McIDAS **S. Limaye/University of Wisconsin, Madison**
11:00	Experimenter's Laboratory for Visualized Interactive Science **E. Hansen/University of Colorado**
11:30	SAVS: A Space Analysis and Visualization System **E. Szuszczewicz/Science Applications International Corporation**
12:00	LUNCH and DEMONSTRATIONS
1:30 p.m.	VIS5D and VISAD Visualization **W. Hibbard/University of Wisconsin**
2:00	LinkWinds: A Distributed System for Visualizing and Analyzing Multivariate and Multidisciplinary Data **A. Jacobson/Jet Propulsion Laboratory**
2:30	DataHub: Knowledge-based Assistance for Science Visualization and Analysis Using Large Distributed Databases **T. Handley, Jr./Jet Propulsion Laboratory**
3:00	BREAK and DEMONSTRATIONS
3:30	Visualization Discussion: Special Features Needed by Earth and Space Science and How to Combine Features for More Powerful Programs **M. Botts/University of Alabama**
5:00	Adjourn to Reception
6:15	**Reception at NCAR**

Thursday, August 13, 1992

8:00 a.m. Coffee

PROGRAMMATIC ISSUES AND ANSWERS SESSION

Chaired by G. Mucklow

8:30 Opening Remarks
G. Mucklow/NASA Office of Space Science and Applications

9:00 NASA Science Internet (NSI) Developments
C. Falsetti/Ames Research Center

9:30 Center of Excellence in Space Data and Information Sciences (CESDIS) Role
R. Miller/CESDIS

10:00 BREAK and DEMONSTRATIONS

10:30 Science Discipline Discussion: Bottlenecks and Problem Areas in Earth and Space Science Processes
Chaired by A. Walton
Astrophysics: T. Ayres/University of Colorado
Earth Science: M. Botts/University of Alabama, Huntsville
Space Physics: J. Willett/NASA Office of Space Science and Applications
Life Sciences: R. Jackson/Ames Research Center
Solar System Exploration: S. Lee/University of Colorado

12:00 LUNCH and DEMONSTRATIONS

1:30 Opportunities for Future AISRP Research Discussion
G. Mucklow/NASA Office of Space Science and Applications

2:00 Technology Transfer Discussion: Test Beds, Demonstrations, Availability/Distribution, User Experience, Software Support, COSMIC Role, Development of Data Test Suites to Evaluate Tools
T. Handley/Jet Propulsion Laboratory

3:00 BREAK and DEMONSTRATIONS

3:30 Collaboration and Communication Discussion: Interaction of Computer Science Investigators with Earth and Space Science Investigators
G. Mucklow/NASA Office of Space Science and Applications
G. Johnston/NASA Office of Aeronautics and Space Technology
J. Giffin/NASA Office of Aeronautics and Space Technology

4:30 Summary and Action Items
G. Mucklow/NASA Office of Space Science and Applications

5:00 Adjourn

AISRP WORKSHOP II
Revised Agenda for Thursday, August 13, 1992

8:00 a.m. Coffee

PROGRAMMATIC ISSUES AND ANSWERS SESSION
Chaired by G. Mucklow

8:30 Opening Remarks
G. Mucklow/NASA Office of Space Science and Applications

8:45 Electronic Mail and News Groups
R. Barth/Goddard Space Flight Center

9:00 NASA Science Internet (NSI) Developments
C. Falsetti/Ames Research Center

9:30 Center of Excellence in Space Data and Information Sciences (CESDIS) Role
R. Miller/CESDIS

10:00 BREAK and DEMONSTRATIONS

10:30 Science Discipline Discussion: Bottlenecks and Problem Areas in Earth and Space Science Processes
Chaired by A. Walton
Astrophysics: **T. Ayres**/University of Colorado
Earth Science: **M. Botts**/University of Alabama at Huntsville
Space Physics: **J. Willett**/NASA Office of Space Science and Applications
Life Sciences: **R. Jackson**/Ames Research Center
Solar System Exploration: **S. Lee**/University of Colorado

12:00 LUNCH and DEMONSTRATIONS

1:00 Technology Transfer Discussion: Test Beds, Demonstrations, Availability/Distribution, User Experience, Software Support, COSMIC Role, Development of Data Test Suites to Evaluate Tools, Proprietary vs. Government Supported Tools
T. Handley/Jet Propulsion Laboratory

2:30 Collaboration and Communication Discussion: Interaction of Computer Science Investigators with Earth and Space Science Investigators
G. Mucklow/NASA Office of Space Science and Applications
G. Johnston/NASA Office of Aeronautics and Space Technology
J. Giffin/NASA Office of Aeronautics and Space Technology

3:00 BREAK and DEMONSTRATIONS

3:30 Opportunities for Future AISRP Research Discussion
G. Mucklow/NASA Office of Space Science and Applications

4:00 Summary and Action Items

4:30 Adjourn

WORKSHOP II

ATTENDEES
APPLIED INFORMATION SYSTEMS RESEARCH PROGRAM (AISRP)

Arrott, Matthew
National Center for
Supercomputing Applications
Beckman Institute for Advanced
Science & Technology
University of Illinois
605 E. Springfield Avenue
Champaign, IL 61820
(217) 244-6829
FAX: (217) 244-1987
E-Mail: arrott@ncsa.uiuc.edu

Ayres, Tom
University of Colorado
Campus Box 389 CASA
Boulder, CO 80309-389
(303) 492-4051
FAX: (303) 492-4052

Barth, C. Wrandle
NASA Science Internet
Hughes STX Corporation
4400 Forbes Boulevard
Lanham, MD 20706
(301) 794-5159
FAX: (301) 306-1010

Berkin, Andrew
Jet Propulsion Laboratory
M.S. 183-501
4800 Oak Grove Drive
Pasadena, CA 91109
(818) 354-0728
FAX: (818) 354-0966
E-Mail:
berkin@krazy.jpl.nasa.gov

Bertini, Alice
Center for Astrophysics and
Space Astronomy
Campus Box 389
University of Colorado
Boulder, CO 80309-6422
(303) 492-6422
FAX: (303) 492-4052
E-Mail:
Bertini@helios.colorado.edu

Blanchard, Paul E.
Science Applications
International Corporation
1710 Goodridge Drive
McLean, VA 22102
(703) 556-7108
FAX: (703) 821-1134
E-Mail: Internet:
blanchard@mclapo.saic.com
SPAN: mclapo::blanchard

Botts, Michael
Marshall Space Flight Center
National Aeronautics and Space
Administration
ES 43
Huntsville, AL 35812
(205) 544-1308
E-Mail:
mibotts@nasamail.nasa.gov
M.Botts/OMNET

Bowman, Kenneth P.
Climate System Research
Program
Texas A & M University
College Station, TX 77843
(409) 862-4060
FAX: (409) 845-6331
E-Mail:
bowman@uiatma.atmos.uiuc.edu

Boyd, William
National Center for
Atmospheric Research
P.O. Box 3000
Boulder, CO 80307
(303) 497-8159
E-Mail: boydw@ncar.ucar.edu

Bredekamp, Joseph H.
Chief Information Systems Branch
Code SMI
National Aeronautics and Space Administration
Washington, DC 20546
(202) 453-1505
FAX: (202) 472-5572
E-Mail:
JBredekamp@nasamail.nasa.gov

Buzbee, William
Director
National Center for Atmospheric Research
P.O. Box 3000
Boulder, CO 80307
E-Mail:
buzbee@bierstadt.scd.ucar.edu

Campbell, Nancy K.
Senior Administrator
Center of Excellence in Space Data & Information Science (CESDIS)
Code 930.5
Goddard Space Flight Center
National Aeronautics and Space Administration
Greenbelt, MD 20771
(301) 286-4099
FAX: (301) 286-5152
E-Mail:
CESDIS@amarna.gsfc.nasa.gov
Campbell@amarna.gsfc.nasa.gov

Carlson, Hallie
NASA Science Internet
Ames Research Center
M.S. 233-8
Moffett Field, CA 94035
(415) 604-6777
FAX: (415) 604-0063
E-Mail:
hallie@nsipo.arc.nasa.gov

Davis, Randy
Laboratory for Atmospheric and Space Physics
University of Colorado
Campus Box 392
Boulder, CO 80309
(303) 492-6867
FAX: (303) 492-6444
E-Mail: AQUILA::DAVIS

Delis, Alexios
Computer Science Department
University of Maryland
College Park, MD 20742
(301) 405-2714
FAX: (301) 405-6707

Djorgovski, George
Department of Astronomy
MS 105-24
California Institute of Technology
Pasadena, CA 91125
(818) 356-4415
FAX: (818) 568-1517
E-Mail:
George@DEIMOS.CALTECH.EDU
also 6035::GEORGE

Doty, Brian E.
University of Maryland
Departmet of Meteorology
College Park, MD 20742
(301) 405-5356
FAX: (301) 314-9482

Emery, William
Colorado Center for Astrodynamics Research
University of Colorado
CB 341
Boulder, CO 80309
(303) 492-8591
(303) 492-2825
E-Mail:
BEmery@nasamail.nasa.gov

Falsetti, Christine
NSI Project Manager
MS 233-8
Ames Research Center
National Aeronautics and Space Administration
Moffett Field, CA 94035-1000
(415) 604-6935
FAX: (415) 604-0063
E-Mail:
falsetti@nsipo.ar.nasa.gov

Freeman, Anthony
Jet Propulsion Laboratory
MS 300-235
4800 Oak Grove Drive
Pasadena, CA 91109
(818) 354-1887
FAX: (818) 393-6443
E-Mail:
Freeman@jplmrs.jpl.nasa.gov

Friedland, Peter
MS 269-2
Ames Research Center
Moffett Field, CA 94024
(415) 604-4277
FAX: (415) 604-3594
E-Mail:
PFriedland@nasamail.nasa.gov

Giffin, Geoffrey A.
Code R
National Aeronautics and Space Administration
Washington, DC 20546
(202) 453-2869
E-Mail: NASA Mail: GGiffin

Goodrich, Charles C.
Astronomy Department
University of Maryland
College Park, MD 20742
(301) 405-1516
FAX: (301) 314-9067
E-Mail: ccg@astro.umd.edu

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

Handley, Jr., Thomas J.
MS 180-401
Jet Propulsion Laboratory
4800 Oak Grove Drive
Pasadena, CA 91009
(818) 354-7009
FAX: (818) 393-6800
E-Mail:
THandley@spacemouse.nasa.gov

Hansen, Elaine
Colorado Space Grant College
University of Colorado
Box 520
Boulder, CO 80309
(303) 492-3141
FAX: (303) 492-2825
E-Mail:
EHansen@nasamail.nasa.gov

Hibbard, William
Space Science and Engineering Center
1225 W. Dayton Street
Madison, WI 53706
(608) 263-4427
FAX: (608) 263-6738
E-Mail:
whibbard@vms.macc.wisc.edu

Hill, John
Houston Advanced Research Center
4800 Research Forest Drive
Houston, TX 77381
(713) 363-7922
FAX: (713) 263-7923

Jackson, Robert W.
Ames Research Center
MS 244-19
National Aeronautics and Space Administration
Moffett Field, CA 94035-1000
(415) 604-5921
FAX: (415) 604-6997
E-Mail: NASA Mail
RWJACKSON.NASA.GOV

Jacobson, Allan
Jet Propulsion Laboratory
Ms 183-501
4800 Oak Grove Drive
Pasadena, CA 91109
(818) 354-0692
FAX: (818) 354-0966
E-Mail: BudJ@apex.nasa.gov

Johnston, Gordon I.
Code RSS
National Aeronautics and Space Administration
Washington, DC 20546
(202) 453-2851
FAX: (202) 426-0001

Jorgensen, Charles C.
Ames Research Center
National Aeronautics and Space Administration
M.S. 269-3
Moffett Field, CA 94035-1000
(415) 604-6725
FAX: (415) 604-6997
E-mail:
Jorgensen@pluto.arc.nasa.gov

Keller, Richard
Sterling Software
AI Research Branch
Ames Research Center
MS 269-2
National Aeronautics and Space Administration
Moffett Field, CA 94035-1000
(415) 604-3388
FAX: (415) 604-3594
E-Mail:
Keller@ptolemy.arc.nasa.gov

Kelly, Timothy
Colorado Center for Astrodynamics Research
University of Colorado
CB 431
Boulder, CO 80309
(303) 492- 8591
FAX: (303) 492-2825

Kinter III, James L.
Department of Meteorology
University of Maryland
College Park, MD 20742
(301) 405-5384
FAX: (301) 314-9482
E-Mail: Kinter@cola.umd.edu

Kipple, Allison
Colorado Space Grant College
University of Colorado
Box 520
Boulder, CO 80309
(303) 492-3141
FAX: (303) 492-2825
E-Mail:
Kipple@aludra.colorado.edu

Klemp, Joseph
National Center for
Atmospheric Research
P.O. Box 3000
Boulder, CO 80307
(303) 497-8902
FAX: (303) 497-8181

Klemp, Margorie K.
LASP Space Technology Building
University of Colorado
Box 590
Boulder, CO 80309
(303) 492-7289
(303) 492-6444
E-Mail:
Margi@rana.colorado.edu

Koch, Steven
Code 912
Goddard Space Flight Center
National Aeronautics and Space Administration
Greenbelt, MD 20771
(301) 286-7188
FAX: (301) 286-4661
E-Mail: MV5::KOCH

Lackman, Robert
National Center for Atmospheric Research
1850 Table Mesa Drive
Boulder, CO 80307-3000
(303) 497-1224
FAX: (303) 497-1137
E-Mail: RLL@ncar.ucar.edu

Lanicci, John M.
United States Air Force
Chief, Meteorological Models
Air Force Global Weather Central
106 Peacekeeper Drive
Suite 2N3
Offutt Air Force Base, NE 68113-4039
(402) 294-4671
FAX: (402) 294-3505
E-Mail: AFGWSYSM@STRATHOST.STRATCOM.AF.MIL

Lawrence, George
Laboratory for Atmospheric and Space Physics
University of Colorado
Campus Box 392
Boulder, CO 80309

Li, Philip
Jet Propulsion Laboratory
M.S. 525-3610
4800 Oak Grove Drive
Pasadena, CA 91109
(818) 306-6035
E-Mail: philip@spacemouse.jpl.nasa.gov

Limaye, Sanjay
Space Science and Engineering Center
University of Wisconsin
1225 W. Dayton Street
Madison, WI 53706
(608) 262-9541
FAX: (608) 262-5974
E-Mail: MADRAF::LIMAYE

Liu, Hua-Kuang
Senior Research Engineer/Group Leader
Jet Propulsion Laboratory
4800 Oak Grove Drive
Pasadena, CA 91109-8099
(818) 354-8935
FAX: (818) 354-9509
E-Mail: HKLiu@neuron1.jpl.nasa.gov

Mankofsky, Alan
Science Applications International Corporation
1710 Goodridge Drive
McLean, VA 22102
(703) 734-5596
FAX: (703) 821-1134
E-Mail: ALAN@MCLAPO.SAIC.COM

McLeod, James J.
San Diego Supercomputer Center
P.O. Box 85608
San Diego, CA 92186-9784
(619) 534-5158
FAX: (619) 534-5113
E-Mail: mcleodj@sdsc.edu

McNabb, David
Department of Astronomy
University of Maryland
College Park, MD 20742
(301) 405-1552
FAX: (301) 314-9067
E-Mail: mcnabb@astro.umd.edu

Miller, Glenn E.
Space Telescope Science
 Institute
3700 San Martin Drive
Baltimore, MD 21218
(410) 338-4738
FAX: (410) 338-1592
E-Mail: Miller@stsci.edu

Miller, Raymond E.
Director
Center of Excellence in
 Space Data & Information
 Science (CESDIS)
Code 930.5
Goddard Space Flight Center
National Aeronautics and Space
 Administration
Washington, DC 20546
(301) 286-4099
FAX: (301) 286-5152
E-Mail: Miller@cs.umd.edu

Mucklow, Glenn H.
Code SMI
National Aeronautics and Space
 Administration
Washington, DC 20546
(202) 453-1678
FAX: (202) 472-5572
E-Mail:
GMucklow@leda.hq.nasa.gov

Parker, John V.
Marshall Space Flight Center
National Aeronautics and Space
 Administration
M.S. E543
Marshall Space Flight Center,
AL 35812
(205) 544-1526
FAX: (205) 544-5760
E-Mail:
jparker@pcjup.msfc.nasa.gov

Preheim, Larry
Jet Propulsion Laboratory
4800 Oak Grove Drive
Pasadena, CA 91109
(818) 306-6042
FAX: (818) 306-6929
NASA Mail: LPREHEIM
ARPANET:PREHEIM@JPLLSI.JPL.NAS
A.GOV

Rotar, Paul
National Center for
 Atmospheric Research
P.O. Box 3000
Boulder, CO 80307
(303) 497-1277
FAX: (303) 497-1137
E-Mail: rotar@ncar.ucar.edu

Searight, Keith
University of Illinois
105 S. Gregory Avenue
Urbana, IL 61801
(217) 333-8132
FAX: (217) 244-4393
E-Mail:
keith@vista.atmos.uiuc.edu

Shaw, Crystal
National Center for
 Supercomputing Applications
405 N. Mathews
5249 Beckman Institute
Urbana, IL 61801
(217) 244-1982
FAX: (217) 244-2909
E-Mail: shaw@ncsa.uiuc.edu

Stetina, Fran
Goddard Space Flight Center
Code 930.8
International Data Systems
 Office
National Aeronautics and Space
 Administration
Greenbelt, MD 20771
(301) 286-5717
FAX: (301) 286-1635

Storer, James A.
Computer Science Department
Brandeis University
Waltham, MA 02554
(617) 736-2714
FAX: (617) 736-2741
E-Mail: 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
E-Mail: szusz@mclapo.saic.com
MCL::SZUSZ

Thompson, David E.
Ames Research Center
Code FIA
MS 269-2
AI Research Branch
Moffett Field, CA 94035-1000
(415) 604-4759
FAX: (415 604-3594
E-Mail:
det@ptolomy.arc.nasa.gov

Walton, Amy
Manager, Science Data Analysis & Computing Systems
Jet Propulsion Laboratory
MS 180-404
4800 Oak Grove Drive
Pasadena, CA 91109
(818) 354-3469
FAX: (818) 393-4468
E-Mail:
AWalton@nasamail.nasa.gov

Willett, James
Code SS
National Aeronautics and Space Administration
Washington, DC 20546
(202) 358-0894

APPENDIX C:

Demonstrations

APPLIED INFORMATION SYSTEMS RESEARCH PROGRAM (AISRP) WORKSHOP II

Laboratory for Atmospheric and Space Physics
Space Technology Research Building
University of Colorado
Research Park
Boulder, CO

AUGUST 11-13, 1992

DEMONSTRATION SCHEDULE

ıgust 11, 1992

10:00	Astrophysics demo DECstation	Alice Bertini, U of Colorado
	Polypaint: Experimenter's Lab for Visualized Interactive Science SGI Indigo	Bill Boyd, NCAR Allison Kipple, U of Colorado
	Distributed Analysis and Visualization Systems for Model and Observational Data SGI Crimson	Steve Koch, GSFC John Hagedorn, GSFC Matt Arrott, NCSA
12:00	Tool for Planetary Atmospheric Modeling SPARCstaton	Rich Keller, Ames Research
	GRADS Grid Analysis and Display System DECstation	Brian Doty, U of Maryland
	VIS5D and VISAD SGI Crimson	William Hibbard, U of Wisconsin
3:00	SAVS: Space Analysis and Visualization System DECstation IBM RS 6000	Alan Mankofsky, SAIC Charles Goodrich, U of Maryland
	Land-Surface Testbed for EOSDIS DECstation	Tim Kelly, U of Colorado
5:00	Interactive Environment for Analysis of Large Earth Observation and Model Data Sets IBM RS-6000	Keith Searight, U of Illinois
	Linkwinds: Distributed System for Visualizing and Analyzing Multivariate and Multidisciplinary Data SGI Crimson	Alan Jacobson, JPL

	DataHub: Knowledge-based Assistance for Science Visualization and Analysis SGI Indigo	Thomas Handley, JPL

August 12, 1992

10:00	GRADS Grid Analysis and Display System DECstation	Brian Doty, U of Maryland
	Distributed Analysis and Visualization Systems for Model and Observational Data SGI Crimson	Steve Koch, GSFC John Hagedorn, GSFC Matt Arrott, NCSA
	Astrophysics demo DECstation	Alice Bertini, U of Colorado
	DataHub: Knowledge-based Assistance for Science Visualization and Analysis SGI Indigo	Thomas Handley, JPL
12:00	SAVS: Space Analysis and Visualization System DECstation IBM RS 6000	Alan Mankofsky, SAIC Charles Goodrich, U of Maryland
	Polypaint: Experimenter's Lab for Visualized Interactive Science SGI Indigo	Bill Boyd, NCAR Allison Kipple, U of Colorado
	Land-Surface Testbed for EOSDIS DECstation	Tim Kelly, U of Colorado
3:00	Linkwinds: Distributed System for Visualizing and Analyzing Multivariate and Multidisciplinary Data SGI Indigo	Alan Jacobson, JPL
	VIS5D and VISAD SGI Crimson	William Hibbard, U of Wisconsi
	Interactive Environment for Analysis of Large Earth Observation and Model Data Sets IBM RS-6000	Keith Searight, U of Illinois
	Tool for Planetary Atmospheric Modeling SPARCstaton	Rich Keller, Ames Research

August 13, 1992

10:00	Astrophysics demo	Alice Bertini, U of Colorado

	DECstation	
	Tool for Planetary Atmospheric Modeling SPARCstaton	Rich Keller, Ames Research
	Interactive Environment for Analysis of Large Earth Observation and Model Data Sets IBM RS-6000	Keith Searight, U of Illinois
	Polypaint: Experimenter's Lab for Visualized Interactive Science SGI Indigo	Bill Boyd, NCAR Allison Kipple, U of Colorado
	VIS5D and VISAD SGI Crimson	William Hibbard, U of Wisconsin
12:00	GRADS Grid Analysis and Display System DECstation	Brian Doty, U of Maryland
	Linkwinds: Distributed System for Visualizing and Analyzing Multivariate and Multidisciplinary Data SGI Indigo	Alan Jacobson, JPL
	Land-Surface Testbed for EOSDIS DECstation	Tim Kelly, U of Colorado
3:00	SAVS: Space Analysis and Visualization System DECstation IBM RS 6000	Alan Mankofsky, SAIC Charles Goodrich, U of Maryland
	DataHub: Knowledge-based Assistance for Science Visualization and Analysis SGI Indigo	Thomas Handley, JPL
	Distributed Analysis and Visualization Systems for Model and Observational Data SGI Crimson	Steve Koch, GSFC John Hagedorn, GSFC Matt Arrott, NCSA

APPENDIX D:

Abstracts

Omit
TO
END

APPENDIX E:

Presentation Material

The following material was presented or distributed at the meeting.
Full size reproductions of this material are available from
Mr. Glenn H. Mucklow, Code SMI, National Aeronautics and
Space Administration, Washington, DC 20546.

MULTIVARIATE STATISTICAL ANALYSIS SOFTWARE TECHNOLOGIES FOR ASTROPHYSICAL RESEARCH INVOLVING LARGE DATA BASES

Dr. S. G. Djorgovski
California Institute of Technology

August 11, 1992

PI: S. G. Djorgovski

Division of Physics, Mathematics, and Astronomy
California Institute of Technology

Collaborators:

W. N. Weir (graduate student, Caltech)
Drs. R. Doyle and U. Fayyad (JPL AI group)

* * * * *

We are conducting a dual effort:

1. Development of a simple, efficient, user-friendly, interactive package for multivariate statistical analysis of relatively small data sets. Development name: STATPROG.

2. Development of a large, complex system to help process and analyse large amounts of data (about 3 Terabytes) from the Digitized Second Palomar Sky Survey (POSS-2). Development name: FRITZ.

STATPROG Summary:

The goal:

Provide a simple, user-friendly toolbox for *multivariate* statistical analysis of relatively small data sets (typically less than 1000 data vectors, and less than 20 dimensions/parameters).

The approach:

A set of programs written in Fortran-77, now running under the VMS OS (soon to be ported to Sun unix), operating on ASCII data files with a simple format: one record per data vector, one column for each parameter, fixed flag value for bad or missing data, recognized column headings, arbitrary number of leading header records.

We use as much of the existing software as possible, e.g., routines described in a monograph *Multivariate Data Analysis* by Murtagh & Heck, *Numerical Recipes*, *Gaussfit* package, various published algorithms for least squares fitting, etc. Some local algorithm development, top-level coding, tests and comparisons.

Status:

Some 15 standalone programs exist and work, and several more are under development or testing. About 10000 lines of code have been produces. Most programs have a common feel, but are not perfectly uniform. Subroutine libraries exist with many useful modules in them.

The package has been exported to two external sites for independent testing by our friends and colleagues.

... STATPROG Summary ...

Initial tests and scientific applications:

This stuff *works* and it produces good science! Several papers based on the use of the package have been published or are in press in major journals, plus a few in various conference contributions or abstracts. We are very pleased with it. These "tests by fire" helped us chase a few bugs out, and smooth some interfaces; a few known bugs remain and are being pursued.

The remaining tasks:

Mostly documentation: better and more extensive code comments, a users' manual and a cookbook, possibly a paper in a technical journal (*PASP* or *Computers in Physics*).

We intend to deposit the whole package + documents in an anonymous ftp account for pickup by any interested parties, or whatever other mechanism is considered appropriate by our sponsors.

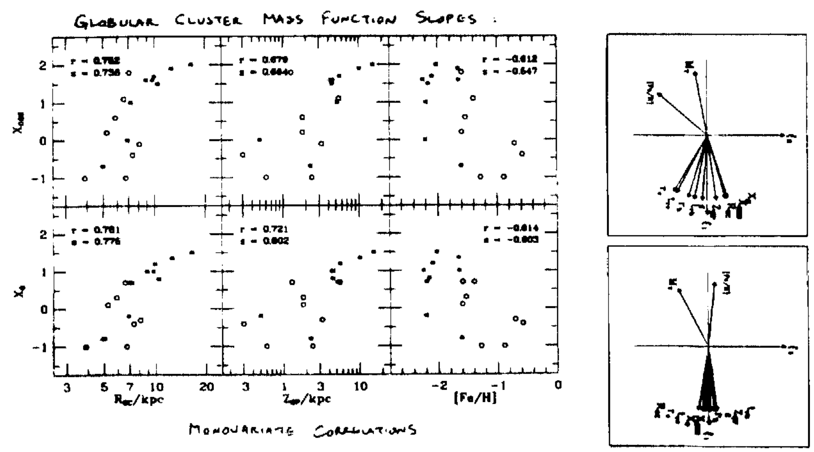
```
Page 1   ... STATPROG ...

$! STATPROG Initialization file                      Last updated: 10 Aug 92
$!
$! Statistical software for use with formatted text data lists of the form
$! described in AAAREAD ME in the STATPROG directory
$!
$
$ PCA         == R STATPROG:PCA                     ! Principal component
$                                                   !   analysis
$
$ NDRACE      :== R STATPROG:NDRACE                 ! Multivariate reduction
$                                                   !   analysis
$
$ BICORR      :== R STATPROG:BICORR                 ! Computes corr. coef  for lin
$                                                   !   comb. of 2 var's vs  1
$
$ THREECORR   :== R STATPROG:THREECORR              ! Computes corr. coef  for lin
$                                                   !   comb  of 3 var's vs  1
$
$ CMATRIX     :== R STATPROG CMATRIX                ! Computes corr. coef  matrix
$                                                   !   for n variables
$
$ CMATRIX2    == R STATPROG CMATRIX2                ! Computes corr  coef  matrix
$                                                   !   for n variables
$
$ LS2FIT      == R STATPROG LS2FIT                  ! Linear regression w/ errors
$                                                   !   (6 methods)
$
$ LS2FIT2     == R STATPROG LS2FIT2                 ! Linear regression w/o errors
$                                                   !   (5 methods)
$
$ GAUSS       == R STATPROG GAUSS/HOUSHOG           ! General n variable regressio
$                                                   !   w/ errors
$
$ GAUSSLIST   == R STATPROG GAUSSLIST               ! Prepare data list for use by
$                                                   !   GAUSS
$
$ XYLIST      == R STATPROG XYLIST                  ! Produce x,y data list from
$                                                   !   lin  comb  of many var's
$
$ MODLIST     == R STATPROG MODLIST                 ! Modify data list
$
$ STATLIST    == R STATPROG STATLIST                ! Compute statistics for
$                                                   !   columns in data list
$
$ OLDGAUSS    == R STATPROG OLDGAUSS                ! Older version of GAUSS
$                                                   !   program above
$
$ CVTLIST     == DIR CVTLIST FOR                    ! General shell for data list
$                                                   !   conversion program
$
$ LS3FIT      == DIR LS3FIT FOR                     ! Unfinished program for lin
$                                                   !   regr. of (X1+aX2) v. Y
```

An example of how one may use the STATPROG package:

- Assemble the data file.
- Run CMATRIX% to generate the matrices of correlation coefficients. Inspect them to find possible interesting correlations, variables.
- Run PCA on different subsets of the input variables. Decide on the dimensionality of the problem. Generate and inspect the correlation vector diagrams to find possible interesting bivariate or trivariate correlations.
- Run BICORR and THREECORR to find the optimal bivariate and threevariate combinations of input variables.
- Run XYLIST to combine the variables into the optimal mixes, per results from BICORR or THREECORR.
- Run LS2FIT and/or LS2FIT2 to compute the slopes and the intercepts for the optimized correlations. Plot and inspect the residuals, etc.
- Alternatively, run GAUSSFIT instead; compare different least-squares fitting methods.
- Repeat for subsets of data as necessary, explore the selection effects, etc.
- Draw up the scientific conclusions, write a paper, and publish!

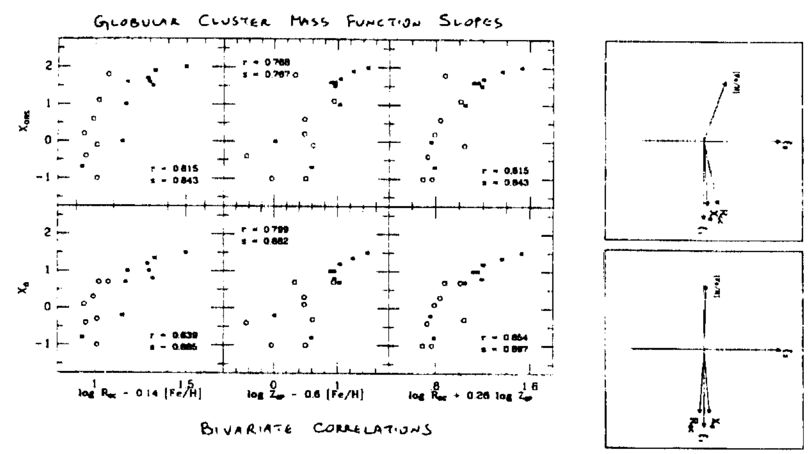
GLOBULAR CLUSTER MASS FUNCTION SLOPES :



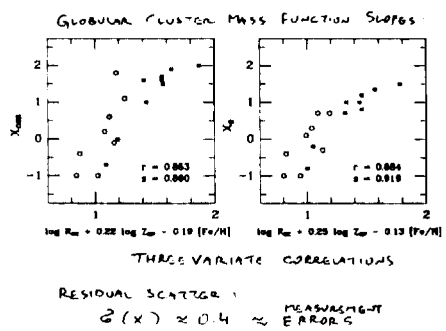
MONOVARIATE CORRELATIONS

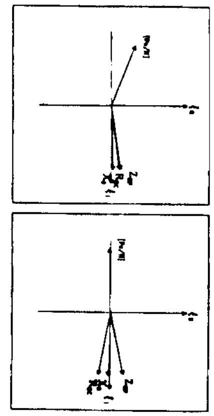
AN EXAMPLE OF A REAL-LIFE SCIENTIFIC APPLICATION OF STATPROG

GLOBULAR CLUSTER MASS FUNCTION SLOPES



BIVARIATE CORRELATIONS





```
Input file: SLOPES DAT
Method: 3      Flag value =     9999 00
Use columns: X0 ( 7)  [Fe/H] ( 8)  Rgc ( 9)  Zgp (10)
Data lines =  17

VARIABLE MEANS AND STANDARD DEVIATIONS FOLLOW.

   Val      X0     [Fe/H]    Rgc      Zgp
 ------  ------  ------  ------  ------
   Mean   0 332   -1 612   0 885   0 424
 StdDev   0 796    0 504   0 149   0 445

PEARSON CORRELATION MATRIX FOLLOWS

    X0    |    1 0000
  [Fe/H]  |   -0 6141    1 0000
   Rgc    |    0 7808   -0 4307    1 0000
   Zgp    |    0 7208   -0 4192    0 5544    1 0000
             --------  --------  --------  --------
               X0       [Fe/H]     Rgc       Zgp

EIGENVALUES FOLLOW

Eigenvalues          As Percentages     Cumul  Percentages
-----------          --------------     ------------------
     2.7825             69 5628               69 5628
     0.6195             15 4870               85 0498
     0.4489             11 2213               96 2711
     0.1492              3 7289              100 0000

EIGENVECTORS FOLLOW

  VBLE.   EV-1     EV-2    EV-3     EV-4    EV-5    EV-6    EV-7
 ------  ------  ------  ------  ------  ------  ------  ------
   X0    -0.5669  0 0712 -0 0860  0 8162
 [Fe/H]   0.4318  0 0737 -0 0499  0.2185
  Rgc    -0.5043  0 3235 -0 6633 -0 4484
  Zgp    -0 4877  0 3563  0 7417 -0 2917

PROJECTIONS OF COLUMN-POINTS FOLLOW

  VBLE.  PROJ-1  PROJ-2  PROJ-3  PROJ-4  PROJ-5  PROJ-6  PROJ-7
 ------  ------  ------  ------  ------  ------  ------  ------
   X0    -0 1091  0 0048 -0 0039  0 0122
 [Fe/H]   0.1443  0 0587 -0 0023  0 0033
  Rgc    -0 1682  0 0217 -0 0303 -0 0067
  Zgp    -0 1627  0 0239  0.0339 -0 0044

NORMALIZED PROJECTIONS OF COLUMN-POINTS ON 1-2 AND 1-3 EIGENPLANES FOLLOW

  VBLE.  PROJ-1  PROJ-2 ALPHA12  PROJ-1  PROJ 3 ALPHA13
 ------  ------  ------  -----  ------  ------  ------
   X0    -0.9997  0 0253   88 55 -0 9998 -0.0208 -268 81
 [Fe/H]   0.9261  0 3772  -67 84  0 9999 -0.0158  -90 91
  Rgc    -0.9918  0 1281   82 64 -0 9841 -0 1775 -259 78
  Zgp    -0.9894  0 1455   81 63 -0 9789  0 2042   78 22
```

Linear correlation coefficients

[illegible]

Spearman's rank correlation coefficients

[illegible]

Kendall's rank correlation coefficients

[illegible]

SAMPLE STATPROG OUTPUT FILES

FRITZ Summary:

The immediate goal:

Develop an AI-assisted software system to process and analyse the digital scans of the Second Palomar Sky Survey (POSS-II) plates

The survey will consist of almost 3,000 photographic plates, in three colors, covering the entire northern sky. They are being digitized at STScI with 1 arcsec sampling, with 23,040 × 23,040 pixels, and 2 bytes per pixel; i.e., about 1 GB per plate, and about 3 TB of image data total.

The resulting catalog is expect to contain over 20 million galaxies, and over 200 million stars, as well as over 0.1 million quasars

The approach:

We are developing a complex system, in collaboration with the JPL AI group. It now runs on a Sparcstation II under the Sun unix OS, but may be ported to a faster machine shortly. The system will detect objects, measure their properties, classify them as stars, galaxies, artifacts, etc., and catalog them. It will use the external CCD imaging for calibration of both the photometry and object classification

Plate scans are processed and calibrated individually, using a large script (AUTOPLATE) and the resulting object catalogs added to the master (whole sky) catalog. The entire production line is a system we call SKICAT.

Experiments are conducted on the object classification algorithms: Neural Nets, Decision Trees (GID3*); both supervised and unsupervised classifiers; in the near future, also AUTOCLASS, and others.

FRITZ Status

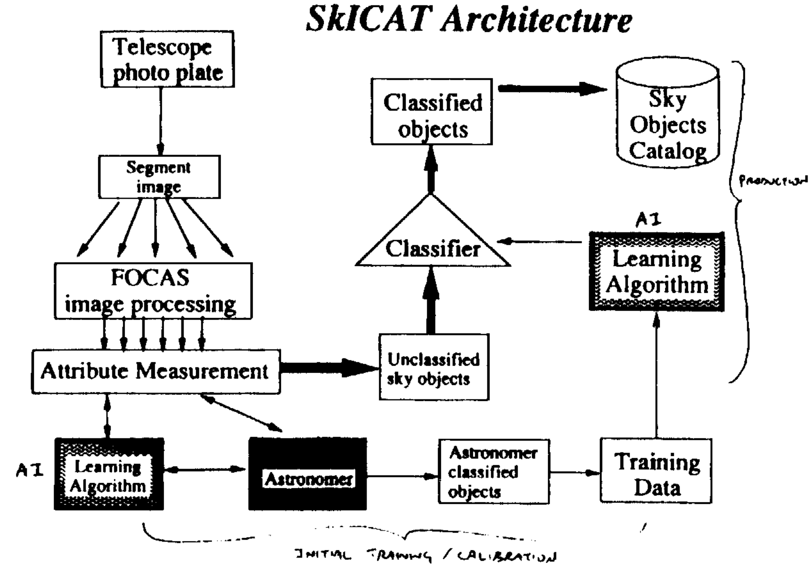
Most of the basic SKICAT system structure exists. We now have a single-plate processing setup (AUTOPLATE) working; what remains to be finished is the plate and CCD frame combining software. We use a modified version of the FOCAS package for object detection and parameter measurements, and the SAS database library for catalog manipulation and analysis.

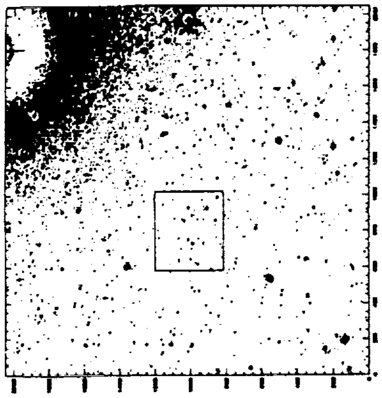
Plates are processed in subsections (footprints) which overlap, both for the sake of efficiency, and in order to deal with the spatial variations in the sky background, PSF shapes, etc. The resulting sub-plate catalogs are combined with lots of internal checking.

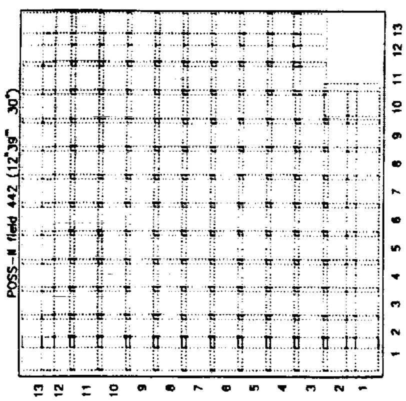
We also have an automatic PSF generating module working, which automatically performs the initial object classification and finds "clean" stellar images. The resulting PSF is then used for a Bayesian resolution classifier of all detected objects.

We have conducted extensive experiments with different classifiers: Neural Nets (multilayer perceptron, with a variety of activation functions and optimization methods), and decision trees (ID3, GID3*). We now favor the later, because of the clear documentability of the classification rules. Both schemes have comparable classification error rates, < 5% for B magnitudes < 20, and < 10% for $B < 21$. This is a full magnitude deeper than in the best works to date, and will enable us to have galaxy clatalogs at least twice as large as the biggest previous efforts.

SkICAT Architecture





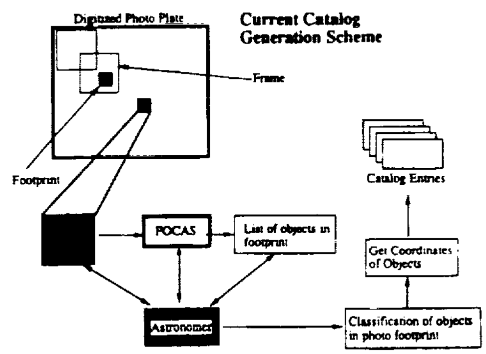


PARAMETERS MEASURED BY FOCAS:

The 18 base-level attributes measured in step 4 are:

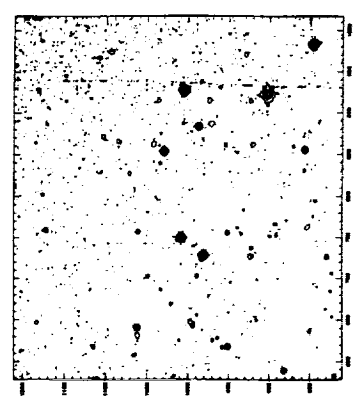
- isophotal magnitude
- isophotal area
- core magnitude
- core luminosity
- sky brightness
- sky sigma (variance)
- image moments (8): ir1, ir2, ir4, r1, r2, ixx, iyy, and ixy.
- eccentricity (ellipticity)
- orientation
- semi-major axis
- semi-minor axis.

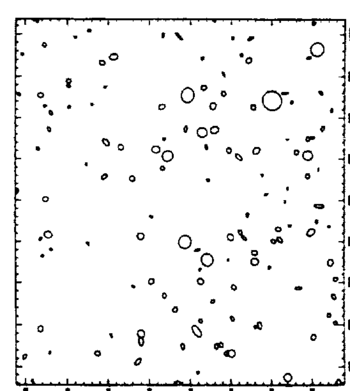
Once all attributes, including the resolution attributes, for each object are measured, step 8 involves performing the final classification for the purposes of the catalog. We are currently considering classifying objects into four major categories: star (s), star with fuzz (sf), galaxy (g), and artifact (long). We may later refine the classification into more classes, however, classification into one of the four classes represents our initial goal.



AUTOPLATE summary: Processing of a single plate scan

- Construct the master plate catalog
- Load image blocks for the current row of footprints
- Mosaic image blocks to form a footprint image
- Assess footprint image quality
- Measure footprint sky and sky sigma
- Intitialize FOCAS footprint catalog
- Detect objects in footprint
- Assess FOCAS sky estimate
- Measure objects in footprint
- Construct re-normalized parameters
- Classify sure-thing stars and construct PSF template
- Apply FOCAS resolution measurement routine
- Save footprint catalog in SAS format
- Match footprint catalog with catalog to left and bottom, checking for consistency
- Go to the next footprint, if there is one
- Check the current row of footprints for trends and outliers among key parameters
- Save the current row's footprint catalogs in the master plate catalog
- Mosaic the current row's sky and reduced images into the master plate images
- Delete the row and go to the next, if there is one
- Check columns of footprints for trends and outliers among key parameters





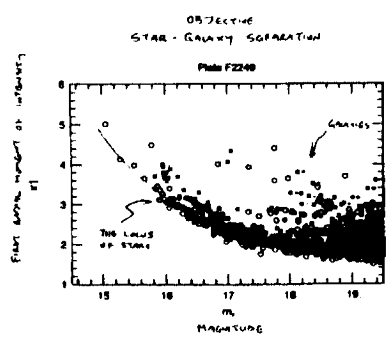


Figure 1: The intensity weighted first moment radius vs. isophotal magnitude for objects from two differently spaced regions on plate F2240. Note that the distribution of point sources (i.e., the 'stellar ridge') changes significantly as a function of position on the plate. It is difficult to design simple and general classification rules based upon such non-uniform attributes.

11

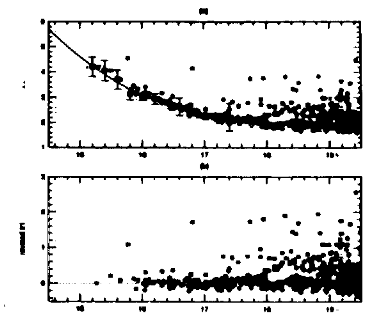
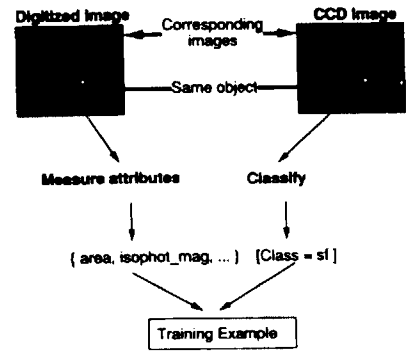


Figure 2: (a) A fourth order polynomial fit to the estimated mode of the FOCAS ir1 parameter distribution in each 0.2" bin; (b) the revised attribute obtained by subtracting the fit to the stellar locus from the original ir1 values. Stars and galaxies are more consistently distributed in this revised parameter space both within and between plates.

12

Classification of faint objects



Why is the accurate object classification important?

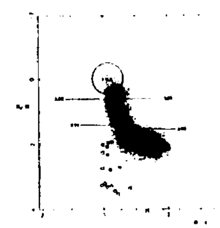
An example of a scientific challenge:

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

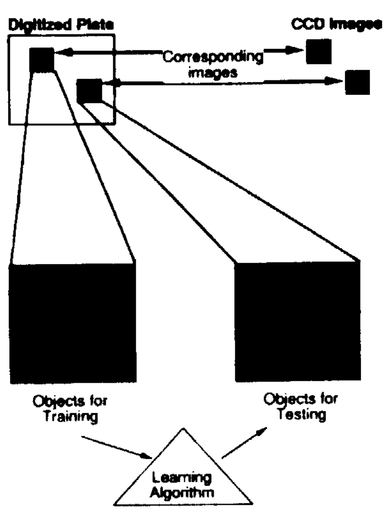
Multicolor selection may be able to bring the number of candidate objects down to about 10,000. However, follow-up spectroscopy with a return of ~ 5% would be prohibitively expensive.

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!



Verification Method

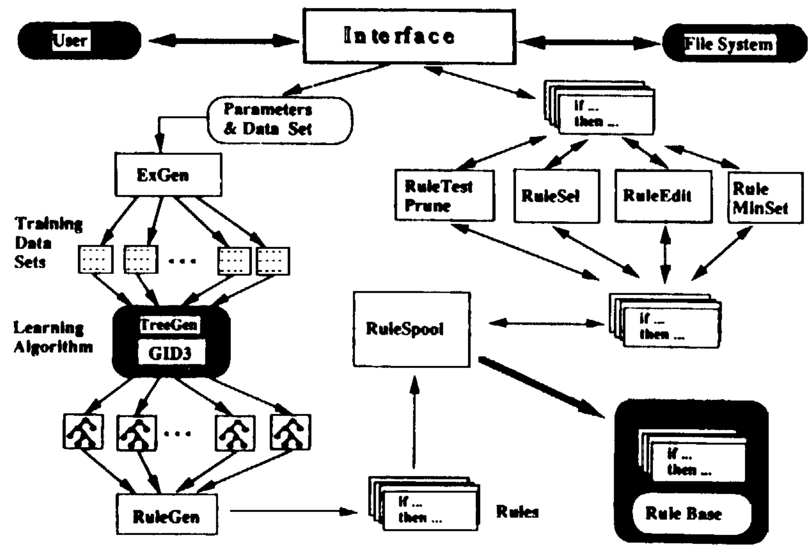


PROCESSING OF A SINGLE PLATE.

The image processing steps that a digitized plate goes through are:

1. Select a frame from the digitized plate.
2. Detection: detect contiguous pixels in the 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: we initially measured 18 *base-level attributes*.
5. Split objects that are "blended" together and re-evaluate attributes.
6. AUTOPSF: select a subset of the objects in the frame and designate them as being "sure-thing stars, form PSF template.
7. Measure *resolution scale* and *resolution fraction* attributes for each object: These are obtained by fitting the object to the template of sure-thing stars formed in step 6.
8. Classify objects in image.

Architechture of RULER system.



ORIGINAL PAGE IS
OF POOR QUALITY

Table 1

Test error rates for F2249 & F2268, m < 19.5

	U-S			G-S			S--S		
Learning Algorithm	Avg	Min	Max	Avg	Min	Max	Avg	Min	Max
Decision Tree (O-BTree)	3.35%	1.43%	5.69%	4.87%	3.00%	6.70%	5.95%	3.32%	9.73%
Decision Tree (GID3*)	3.16%	1.30%	4.97%	5.50%	3.20%	8.17%	5.12%	2.45%	7.91%
Neural Network (VMM, 2 Hidden Nodes)	1.73%	0.00%	4.17%	0.95%	0.00%	2.27%	8.40%	3.76%	14.47%
Neural Network (VMM, 4 Hidden Nodes)	2.16%	0.40%	3.97%	0.96%	0.00%	1.97%	6.96%	2.72%	12.92%
Neural Network (BP, 2 Hidden Nodes)	2.83%	0.20%	5.09%	1.56%	0.20%	3.97%	2.91%	0.52%	5.73%

Test error rates for F2249 & F2268, m < 19

	U-S			G-S			S--S		
Learning Algorithm	Avg	Min	Max	Avg	Min	Max	Avg	Min	Max
Decision Tree (O-BTree)	3.71%	1.34%	6.80%	8.03%	4.59%	16.44%	8.03%	4.59%	16.44%
Decision Tree (GID3*)	3.28%	1.19%	5.97%	8.47%	3.78%	16.47%	8.47%	3.78%	16.47%
Neural Network (VMM, 2 Hidden Nodes)	2.71%	0.00%	7.31%	1.31%	0.00%	3.88%	1.31%	0.00%	3.88%
Neural Network (VMM, 4 Hidden Nodes)	2.93%	0.26%	8.13%	1.05%	0.00%	2.82%	1.05%	0.00%	2.82%
Neural Network (BP, 2 Hidden Nodes)	3.44%	0.53%	6.72%	2.02%	0.27%	5.72%	2.02%	0.27%	5.72%

Where U-S = Unkown object classified as star
G-S = Galaxy classified as star
S--S = Star not classified as star

Table 2

Test error rates for F831, m < 19.5

	U-S			G-S			S--S		
Learning Algorithm	Avg	Min	Max	Avg	Min	Max	Avg	Min	Max
Decision Tree (O-BTree)	6.70%	5.00%	8.50%	4.00%	2.00%	6.50%	9.40%	5.00%	13.50%
Decision Tree (GID3*)	6.50%	5.50%	8.00%	3.70%	2.00%	5.00%	9.80%	6.50%	16.00%
Neural Network (VMM, 2 Hidden Nodes)	4.50%	2.00%	8.00%	1.80%	0.50%	3.50%	15.00%	1.50%	25.50%
Neural Network (VMM, 4 Hidden Nodes)	7.50%	6.50%	8.00%	2.60%	1.50%	4.00%	7.60%	3.00%	11.00%
Neural Network (BP, 2 Hidden Nodes)	8.20%	7.50%	9.00%	3.60%	2.00%	5.00%	3.10%	0.00%	7.50%

Test error rates for F831, m < 19

	U-S			G-S			S--S		
Learning Algorithm	Avg	Min	Max	Avg	Min	Max	Avg	Min	Max
Decision Tree (O-BTree)	3.03%	1.36%	6.69%	1.39%	0.35%	3.56%	1.39%	0.35%	3.56%
Decision Tree (GID3*)	3.91%	1.69%	10.75%	1.63%	0.68%	4.06%	1.63%	0.68%	4.06%
Neural Network (VMM, 2 Hidden Nodes)	1.97%	0.35%	5.38%	0.34%	0.00%	0.68%	0.34%	0.00%	0.68%
Neural Network (VMM, 4 Hidden Nodes)	5.12%	2.38%	16.11%	0.74%	0.35%	2.00%	0.74%	0.35%	2.00%
Neural Network (BP, 2 Hidden Nodes)	5.08%	2.38%	14.55%	0.78%	0.35%	1.85%	0.78%	0.35%	1.85%

Where U-S = Unkown object classified as star
G-S = Galaxy classified as star
S--S = Star not classified as star

FRITZ: Next Steps

- Matching and integrating of different plate catalogs and CCD data, automatic recalibrations as the better data come in.
- Further experiments with classifiers: novel NN schemes, additional object parameters, combined-parameter rules, unsupervised classification, etc.
- Experiments with optimized pre-classification (e.g., by redefining the object parameters, clustering analysis, PCA, Autoclass...)
- In the future, as the storage technologies improve, remeasure the objects and define new parameters on the basis of expert system (mark II classifier) suggestions.
- Scientific analysis of the resulting Northern Sky Catalog, including the statistical analysis tools (Super-STATPROG?), catalog manipulation (SAS with a GUI), etc.
- Use for other applications: massive CCD surveys, Earth-grazing asteroid patrol surveys, ...

Long-Term Plans:

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. 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. While we are dealing here with a particular problem of image analysis, the experience gained in this project should be of a more general use.

Put simply, we are proposing to *generalize the concept of an astronomical catalog*, from that of a fixed, or sometimes updated list of objects, to *a dynamical, evolving entity*, which can be constantly improved as better information becomes available. This process of an evolving catalog naturally leads to a better understanding of the input data and their limits, and, as a byproduct, provides constant and uniform quality control.

The system we are developing will be a tool, a research aid which would perform moderately sophisticated, but highly repetitive and tedious tasks; it will not be a replacement for a human astronomer. The essence of this effort is to introduce expert systems and machine learning techniques to the analysis of modern astronomical data in a major way. We see our effort as a development of software tools for the astronomy of the turn of the century, and beyond.

MULTI-LAYER HOLOGRAPHIC BIFURCATIVE NEURAL NETWORK FOR REAL-TIME ADAPTIVE EOS DATA ANALYSIS

Dr. Hua-Kuang Liu
Jet Propulsion Laboratory

August 11, 1992

JPL
Jet Propulsion Laboratory
California Institute of Technology

Multi-Layer Holographic Bifurcative Neural Network
For Real-Time Adaptive EOS Data Analysis

Hua-Kuang Liu
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California 91109

AISRP WORKSHOP II
Boulder, Colorado
August 11, 1992

OUTLINE

I. Introduction

II. Theoretical Discussion
  2.1 Basic Model
  2.2 Principle of Bifurcation

III. Experiment
  3.1 Set-up
  3.2 Results

IV. Conclusion

A bifurcating optical pattern recognizer (BIOPAR)

* **Physics**
  - a nonlinear gain saturation memory medium
  - scattered coherent light waves
  - microscopic holographic gratings
  - neuromorphic

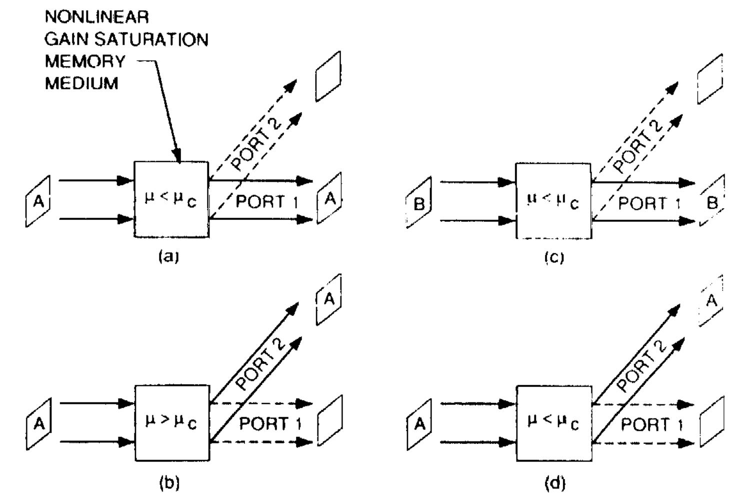
* **Training**
  - Input image intensity greater than a threshold
  - Generation of microscopic waves by scattering of the input wave from the doped impurities in the medium
  - Holographic interference patterns formed as refractive index gratings due to migration and trapping of the photon-induced electrons
  - Microscopic synaptic interconnections

* After training is completed, any new input applied to the medium can be detected at the output in two distinctly different directions which are dependent on the content of the input.

* If the new input is completely different from the memorized input, the new input wave will follow its original path affected only by the refraction of the crystal.

* If the new input is the same as the memorized input, it will be diffracted to a new direction. The intensity and quality of the diffracted image signifies the similarity between the new and the memorized input.

* A neuromorphic model of the BIOPAR with experimental demonstration using a barium titanate crystal are presented.



II. THEORETICAL DISCUSSION

* The commonality between the neural net computation and the nonlinear optical media wave interaction is that they both are governed by dispersive nonlinear wave equations.

* In terms of bio-physics, the membrane potential of each neuron is a nonlinear time-delayed function of the summation of synaptic weighted inputs from all connected cells. Similarly, based on the wave interaction principle in photorefractive crystals, the field at any lattice site in a nonlinear optical medium can be written as a function of the delayed fields at neighboring sites.

* **Assumptions**

  Existence of an excitability threshold expressed in terms of energy or potential

  Input energy to a neuron below the threshold, neuron state stable and quiescent

  Threshold exceeded, critical modes in the form of standing or travelling waves bifurcate(Hopf bifurcation)

2.1 Basic Model

* The neural network with nearest neighbor local interconnections is analogous to coupled waves in a narrow spectrum about a center frequency ω_c in a nonlinear optical medium with gain, saturation, and memory.

* When the input intensity is higher than a threshold μ_c, bifurcating waves are generated.

* These waves oscillate at a fast carrier frequency ω_c and are modulated by a slowly varying envelope $v(\mathbf{r},t_s)$, a function of space coordinates $\mathbf{r}$ and a suitably defined slow time variable t_s.

One solution of the wave function

$$\Psi(\mathbf{r},t,t_s) = v(\mathbf{r},t_s)\exp(j\omega_c t) + v^*(\mathbf{r},t_s)\exp(-j\omega_c t), \qquad (1)$$

"*" : complex conjugate.

By expanding μ in a small neighborhood of μ_c and add a memory nonlinearity to the system and letting $v(\mathbf{r},t_s)\equiv v$

$$j\epsilon^2(\delta k^2/\delta\omega)_c(\delta v/\delta t_s) = \nabla^2 v + [k^2(\mu,j\omega_c) + D - h|v|^2]v, \qquad (2)$$

where $k^2\equiv k^2(\mu,j\omega)$, and the cubic term "$-h|v|^2v$" with h as a constant is the lowest-order nonlinear term. Solutions of Eq. (2) yield properties of wave functions.

Assume k^2 varies linearly versus $|v|^2$ and

$$dk^2/dt_s = c|v|^2 - bD, \qquad (3)$$

where b and c are constants, D is delayed nonlinear memory.

The delayed nonlinear memory D is produced by the change in membrane and synaptic conductance properties with continued excitation. In the Hopf expansion, ϵ^2 is a parameter derived from the small amplitude solutions.

$$D = c\exp(-bt_s)\int_{-\infty}[\exp(bt')]|v(t')|^2dt', \qquad (4)$$

with the upper limit of the integration set at t_s.

An initial condition $v(\mathbf{r},0)$ corresponds to the sensory input of a real image at t=0 to a nonlinear PRC.

2.2 Principle of Bifurcation

A perturbation expansion of small amplitude solutions centered in a small neighborhood of μ_c with a series of functions v_1, v_2, and v_3 etc.

$$v = \epsilon v_1 + \epsilon^2 v_2 + \epsilon^3 v_3 + \ldots \qquad (5)$$

Using the first order term, a general solution as summation of waves of amplitude A_j and phase ϕ_j, with j = 1 ..., to N :

$$v_1(r,t_s) = \Sigma\, A_j(t_s)\phi_j(r), \qquad (6)$$

where $\phi_j(r) \sim \exp(jk_j \bullet r)$.

Modulation depth may be approximately written as

$$k_{1m} = C\, A_{s1}A_{sm}^{*}. \qquad (9)$$

Matrix element in Hebb's outer-product learning!

Explicitly, this "neural synaptic" matrix may be expressed by

$$\{k\} = C\, v_s v_s^{*}. \qquad (10)$$

The superscript "*" represents transpose operation and c and C are constants.

(a) Training via the Presentation of 2-D Patterns

Input image $I_s(r)$ initialized modal superposition

$$v_s(0) = \Sigma\, A_{sj}(0)\phi_j \qquad (7)$$

which enables self-organization of holographic gratings with a set of grating vectors $\mathbf{k}_{1m}=(\mathbf{k}_1-\mathbf{k}_m)$ and modulation depth $k_{1m}(t_s)$:

$$k_{1m} = c\,\exp(-bt_s) \int \exp(b\tau)\, A_{s1}(\tau)A_{sm}^{*}(\tau)\, d\tau. \qquad (8)$$

(b) Memory via Holographic Grating Formation

* After the holographic gratings representing the synaptic interconnections governed by Eq.(10) are written and stored in the medium, the training is completed.
* A multiple number of images may be learned and memorized in the medium by exposing the medium to the inputs sequentially.
* **<u>No external reference is applied.</u>** Different from conventional holographic recording.

(c) Pattern Recognition via Gainful Associative Retrieval

After the medium learned at least one image, a new image may be applied. The directions of the bifurcating waves are determined by the Bragg conditions obtained in solving Eq.(5) for the third order term v_3. The Bragg conditions of waves diffracted from the memorized gratings are given as:

$$\mathbf{k}_l - \mathbf{k}_m + \mathbf{k}_p = \mathbf{k}_j, \qquad (11)$$

with $l,m = 1,\ldots n$, and $p,j = 1,\ldots N$,

where $n<N$ represents the number of possible excited modes within the memory of the previous image.

Corresponding coupled wave equations of excited bifurcating waves:

$$dA_j/dt_e = (\lambda+\Sigma k_{mm})A_j + \Sigma k_{jm}A_m - 2gA_j\Sigma|A_m|^2+g|A_j|^2A_j-k_{jj}A_j, \qquad (12)$$

where the summation runs from 1 to n, λ is a constant, and gain g dependent on the material properties, crystal orientation, and polarization and incident angle of the input beam.

The terms with the gain parameter, g, in Eq.(12) are major factors contributing to the growth of the excited bifurcating waves until saturation is reached.

The bifurcating waves may be used as a quantitative indicator of how much the new input resembles the memorized image.

III. Experiment

3-1 Set-up

A Spectra Physics model 2000 argon ion laser is collimated by the pinhole spatial filter(S. F.) and lens L_1 combination. A liquid crystal television(LCTV) spatial light modulator(SLM) or a photographic transparency is used to apply input images or data to the system. The L_2-pin-hole-filter-L_3 combination is used to enhance the contrast of the output image of the LCTV SLM before it is applied to the nonlinear medium, which in this case is a PRC with a high gain coefficient. A barium titanate($BaTiO_3$) is chosen. The crystal is 5mmx5mmx8mm. A Sony CCD camera detected outputs are displayed on a TV monitor, an oscilloscope, or a chart recorder.

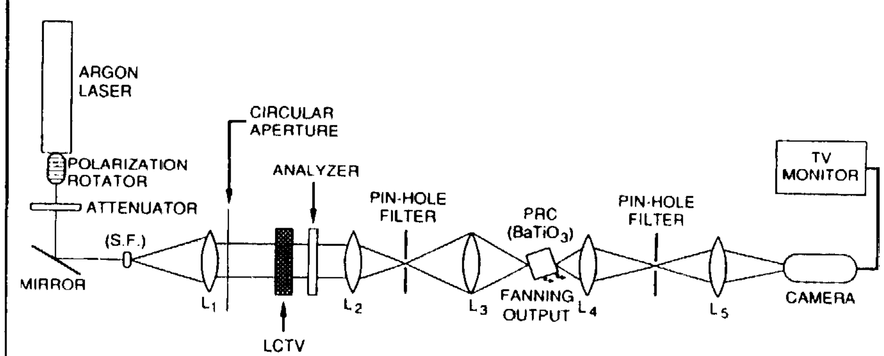
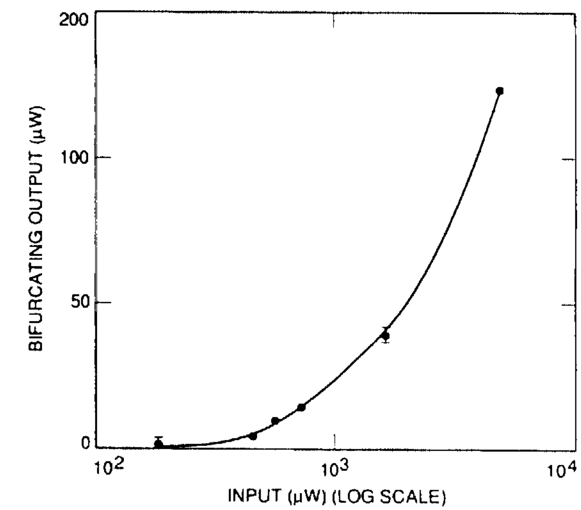


Fig. 1

3-2 Experimental Results

Fig. 3. Before a nonlinear gain saturation memory medium is trained and when an input image "A" with intensity less than the threshold energy, $\mu<\mu_c$, is applied, the output at Port 1 is "A" and the output at Port 2 is zero. This is illustrated in Fig. 3(a). After training with "A" with $\mu>\mu_c$, as illustrated by 3(b), the output at Port 1 becomes zero and the output at Port 2 becomes "A". In Fig. 3(c), it shows that if the medium was trained with input "A" and then applied with an extremely weak "B" of energy $\mu<\mu_c$, the output at Port 1 will be "B" and at Port 2 will be zero. Figure 3(d) shows that if the extremely weak input in Fig. 3(c) is "A" instead of "B", then the output at Port 1 will be zero and the output at Port 2 will be "A".



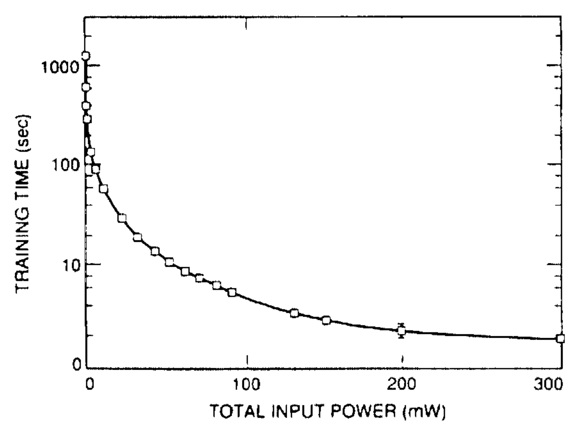


Fig 19

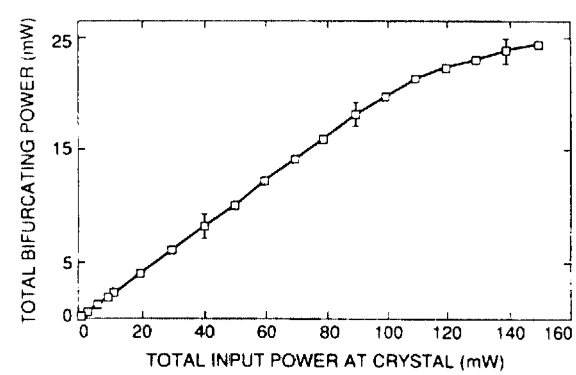
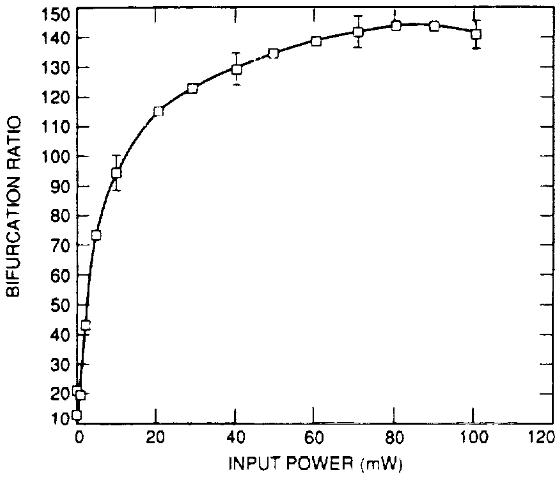


Fig 20



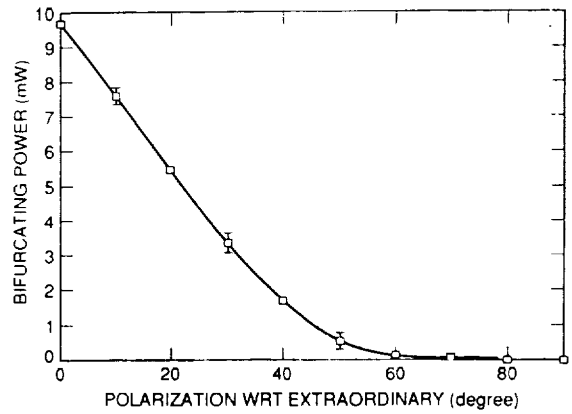


Fig 22

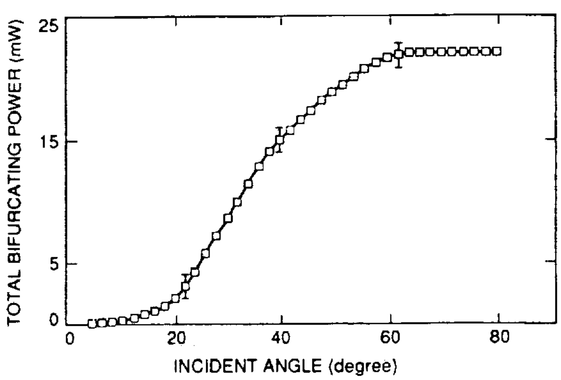


Fig 23

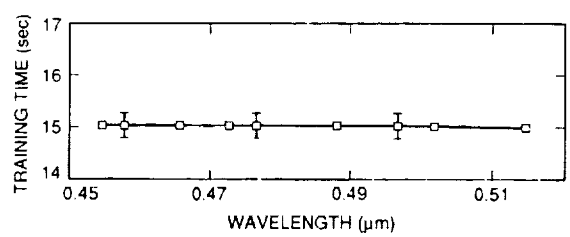


Fig 24

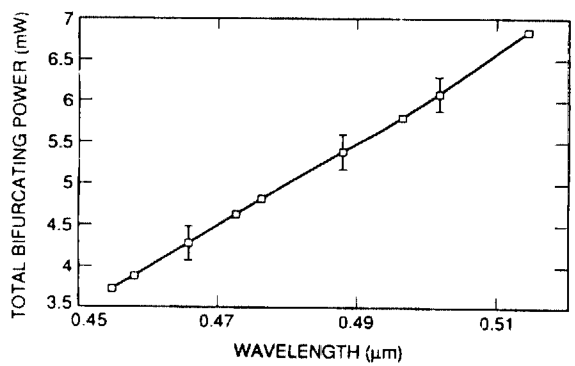


Fig. 25

IV. CONCLUSION

* We have introduced a new class of optical pattern recognition technique which utilizes the bifurcating diffraction phenomenon in nonlinear gain saturation memory media.

* The uniqueness of the BIOPAR is that only a single input beam, instead of one signal and one reference beam, is applied as the input in both the training and addressing process.

* Future research for BIOPAR includes the improvement of the quality and resolution of the bifurcating signals and the application of the BIOPAR for EOS data classification problems.

* **Applications**

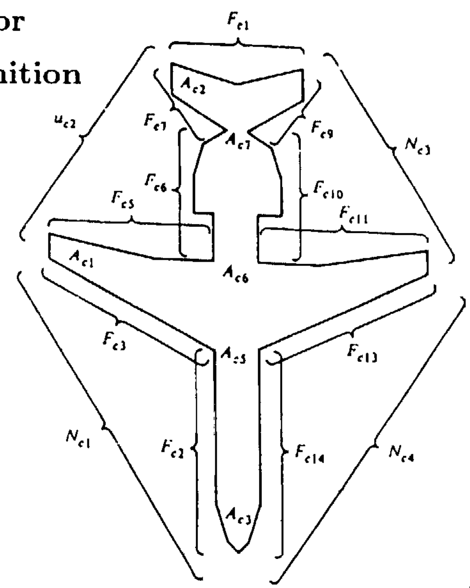
Applications of the Bifurcative Neural Network System

- Optical Pattern Recognition through Syntactic Analysis with Adaptive Neural Learning.
- Optical Arithmetic based on Symbolic Substitution using a Bifurcative Array of Photorefractive Crystal Cells.

Optical Pattern Recognition Application Areas

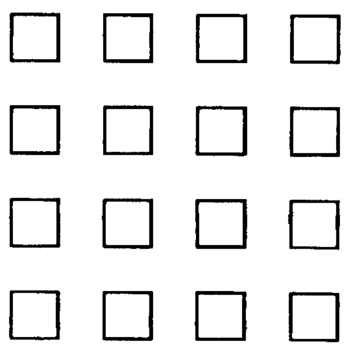
- Cluster analysis for synthetic patterns based on shortest distances.
- Shape analysis of waveforms and contours.
- Character recognition using a syntactic approach.
- Fingerprint recognition using tree grammars.
- Optical image understanding with real-time correlation.

Segmentation of the Airplane BAC111 Contour for Syntactic Shape Recognition



Optical Arithmetic using the Bifurcative Pattern Recognition Array

- Perform Optical Arithmetic using Symbolic Substitution and Signed-Digit Code.
- Constructing a 2-D Array of Bifurcative Pattern Recognition Cells for Fast Symbolic Substitution.
- Exploring Massive Parallelism in Optics with an Adaptive Neural Approach.



A 2-D Array of $BaTiO_3$ Cells for Bifurcative Symbolic Substitution in Performing Optical Arithmetic

Major Technical Achievements

* Developed a new class of brain-in-a-cube-type bifurcating opto-electronic pattern recognition(BOPAR) technique for data classification using high-gain nonlinear photorefractive crystals with an experimental demonstration using barium titanite.

* Developed a new class of shift, rotation, and scale invariant adaptive optical pattern recognition(AOPAR) system for discerning very faint object merged in highly cluttered background and experimentally demonstrated using high-gain photorefractive barium-titanite crystals.

* Developed and demonstrated an unipolar terminal-attractor based associative memory(UTABAM) system via adaptive threshold enabling perfectly correct convergence in the associative recall of inputs where the number of stored states are comparable to the number of neurons thus showing a significant improvement over the Hopfield neural net.

* Discovered a new pyramidal multi-layer multi-resolution optoelectronic data classification system utilizing the perfectly convergent unipolar terminal-attractor based associative memory and the futuristic SEED devices.

NEW TECHNOLOGY INNOVATIONS AND PUBLICATIONS

The following research results are credited to the projects synergistically supported by DARPA and NASA-NRA.

1. A U. S. patent No. 5,005,954 entitled "Method and Apparatus for Second-Rank Tensor Generation" was received by Hua-Kuang Liu. This invention is useful for generations of optical interconnection matrices in real time and is important for future optical computer networking applications.

2. NASA Tech Brief NPO-19098 entitled "Photorefractive Crystal Compresses Dynamic Range of Images" was received by Hua-Kuang Liu. The article received over 150 inquires from U. S. Companies and is listed as one of the most frequently requested among the published Tech Briefs.

3. Three papers authored and co-authored by Hua-Kuang Liu were submitted for publication and presentation at the 1992 Conference on Lasers and Electro-optics at Annaheim, California. The Conference is co-sponsored by IEEE and OSA. The titles of the papers are:

"An Adaptive Invariant Optical Pattern Recognizer"

"Optimizing CGH realization in Electrically-controlled SLMs for Reconfigurable Optical Interconnections and Real-time Zoom Lenses"(with S. Zhou and C. Tan)

" A Comparative Analysis Between The Fourier and Fresnel Digital Holography for 3-D Data Visualization"(with S. Zhou)

4. Three papers were authored or co-authored by Hua-Kuang Liu were presented at the 1991 Annual Meeting of the Optical Society of America at San Jose, Ca. during November 4-9, 1991. The papers are entitled:

" Birfucative Optical Pattern Recognition"
" Holographic Lens Design Using the Contact-Screen Method"(with F. Peng, A. A. S. Awwal, and M. A. Karim)
" Adaptive Reconfigurable Interconnects Using a LCTV" (with A. A. S. Awwal, Alastair D. McAulay, and Junqing Wang)

Acknowledgement

The presented research was jointly sponsored by the Defense Advanced Projects Agency and the Information Systems Branch, Office of Space Science and Applications, National Aeronautics and Space Administration, and was performed by the Center for Space Microelectronics Technology, Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

The experimental assistance of J. Diep is appreciated.

Major Collaborators: Dr. Kai Hwang(USC)
Dr. Pochi Yeh, Dr. Shaomin Zhou(UCSB)
Dr. John Wu(Auburn Univ.)

DEVELOPMENT OF AN EXPERT DATA REDUCTION ASSISTANT

Dr. Glenn Miller
Space Telescope Science Institute

August 11, 1992

SPACE TELESCOPE SCIENCE INSTITUTE | AISRP Workshop II, Glenn Miller, August 1992

Development of an Expert Data Reduction Assistant

Glenn Miller, PI

Co-Investigators: Felix Yen, Mark Johnston and Robert Hanisch

Space Telescope Science Institute
3700 San Martin Dr.
Baltimore, MD 21218

Internet: miller@stsci.edu

SPACE TELESCOPE SCIENCE INSTITUTE | AISRP Workshop II, Glenn Miller, August 1992

Data Reduction

The process of converting raw instrumental output into physical measurements, i.e.,

Steps taken to minimize the influence of data acquisition imperfections on the estimation of the desired astronomical quantity (Gilliland 1992)

In this project we focus on the reduction problem, rather than on data analysis or visualization

SPACE TELESCOPE SCIENCE INSTITUTE | AISRP Workshop II, Glenn Miller, August 1992

Data Reduction/Analysis Systems

- IRAF - Image Reduction and Analysis System (NOAO)
- STSDAS - Space Telescope Science Data Analysis System (STScI)
- MIDAS - Munich Interactive Data Analysis System (ESO)
- IDL - Interactive Data Language
- .....Many other systems

Very successful approach

- widespread distribution of these systems
- systems written for one wavelength extended to serve others
- incorporation of independently developed packages

SPACE TELESCOPE SCIENCE INSTITUTE — AISRP Workshop II, Glenn Miller, August 1992

Philosophy of These Systems

- Modular operators which work on standardized types of data files
- Command Language to execute single commands or scripts of commands (in interactive or batch mode)

Advantages:

- Flexibility for the user: individual commands can be chained (or "pipelined") to construct powerful, customized procedures
- Ease of development: Well-defined methods for adding new modules. Thus many programmers and scientists may independently contribute to the growth of a system.
- Standardization

SPACE TELESCOPE SCIENCE INSTITUTE — AISRP Workshop II, Glenn Miller, August 1992

Disadvantages of Existing Systems

- Learning a system isn't easy
  - Commands can be complex with many parameters and even experts don't know the entire system. Users may have to learn more than one system, especially if they work at different institutions or their work is multi-disciplinary/multi-spectral
- Difficult to capture expert knowledge
  - Manuals, on-line help, local gurus have drawbacks
- Data management problem
  - A few night's observations can result in hundreds of data files which must each pass through many reduction steps. Prone to errors which are very difficult to detect
- Iterative nature of real scientific work
  - Usually must reduce data several times as one learns more of data, calibration files, algorithms, etc.

These problems were echoed in a study of EOSDIS by Botts of U of AL, as forwarded to aisrp-members list by Glenn Mucklow

SPACE TELESCOPE SCIENCE INSTITUTE — AISRP Workshop II, Glenn Miller, August 1992

Expert Assistant

An alternative approach which builds on the foundation of these systems

Goal is to develop a system which will act much as a graduate assistant:

- follows instructions on how to do data reduction
- checks for processing and data quality problems
- works longer hours than you do
- flexible in that existing knowledge of reduction procedures can be easily modified for new cases

Related work towards developing expert assistants can be found in the literature in a variety of domains

SPACE TELESCOPE SCIENCE INSTITUTE — AISRP Workshop II, Glenn Miller, August 1992

Features of the Expert Assistant

- Gather information about the available data (typically from header information in the data files).
- Develop a plan for data reduction based on the user's goals, actual properties of the data and on limitations of available resources (e.g. disk space).
- Translate the plan into explicit reduction commands for a specific data analysis system.
- Monitor the plan and its execution for problems (e.g. missing calibration files) and alert the user.
- Be extensible to incorporate new types of data reduction, new analysis modules and new data analysis systems. The Expert Assistant will provide users with tools for this purpose.
- Present powerful and effective user interface including mouse-and-menu graphics (which is also found in non-expert systems) and natural language interface.

Prototype system for calibration of CCD images was developed by Johnston in 1987

SPACE TELESCOPE SCIENCE INSTITUTE | AISRP Workshop II, Glenn Miller, August 1992

How to Best Involve the Scientific Community?

Lead Users:

- Involve scientists in use and independent evaluation of expert assistant as soon as possible (in addition to scientific input of PI and Co-Is [3 of 4 are astronomers])
- We feel that it is important that this tool be used with real data reduction problems as early as possible.
- Even the initial versions of the Expert Assistant will be sufficiently powerful to pay back the Lead Users for their investment of time.

We interviewed several groups at STScI and the reduction problems were very similar

SPACE TELESCOPE SCIENCE INSTITUTE | AISRP Workshop II, Glenn Miller, August 1992

First Lead Users:

Hubble Space Telescope Medium Deep Survey (MDS). Richard Griffiths, Kavan Ratnatunga and others.

A Key HST project. These are projects which were identified by the astronomical community as high scientific importance and involving a large amount of HST observing time. Data is shared by many astronomers with different interests.

Scientific objectives include: serendipitous discoveries and rare objects, morphology and distribution of faint galaxies, active nuclei of distant galaxies, galactic structure, distant solar-system objects

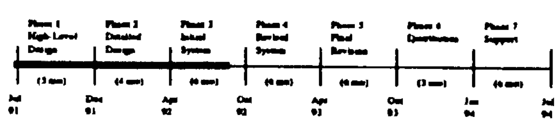
Obtaining CCD image data with Wide Field/Planetary Camera and Faint Object Camera.

Done in parallel with other observations.

Removal of cosmic rays from CCD data is a prime challenge, and the current focus of the Expert Assistant.

SPACE TELESCOPE SCIENCE INSTITUTE | AISRP Workshop II, Glenn Miller, August 1992

Project Schedule and Progress to Date



Have developed initial system:
- language for specifying reduction steps
- expert assistant analyzes existing data
- produces script for data reduction in a specific analysis system

SPACE TELESCOPE SCIENCE INSTITUTE | AISRP Workshop II, Glenn Miller, August 1992

Sample Logfile

```
# LOGIN Tue 17:52:06 04-Aug-92
cl.keeplog = yes
# 17:52:06 clpackage.stsdas - Start (stsdas$stsdas.cl)
stsdas
# 17:52:07 stsdas.wfpc - Start (wfpc$wfpc.cl)
wfpc
combine.logfile = '/marian/data2/mds/draco04Aug92.log'
# 17:52:08 wfpc.combine - Start (wfpc$x_wfpc.e)
combine @/marian/data2/mds/dracoAAAa15095 /marian/data2/mds/dracoAAAa15099 optir
Aug  4 17:52 combine: crreject, lowreject=0., highreject=3.
              Images        N
    w0ul1d03t.c0h[1/4]      1
    w0ul1d02t.c0h[1/4]      1
    w0ul1a03t.c0h[1/4]      1
    w0ul1a02t.c0h[1/4]      1
    w0ul1a01t.c0h[1/4]      1
  --------------------- ------
  /marian/data2/mds/dracoAAAa15099[1/4]
                        ... additional output deleted...
# 18:12:48 wfpc.combine - Stop
# 18:12:49 stsdas.wfpc - Stop
# 18:12:49 clpackage.stsdas - Stop
```

SPACE TELESCOPE SCIENCE INSTITUTE

AISRP Workshop II
Glenn Miller
August 1992

Sample Script

```
#!/bin/sh
# script title: REMOVE-CR-NOISE
# procedure for removing cosmic ray noise
# creation date: 04-Aug-1992, 17:47:04
# create and initialize log file
echo ":-( Draco :-)" > /marian/data2/mds/draco04Aug92.log
# initialize package IRAF
echo cl.logmode = \'errors trace\' >> /marian/data2/mds/dracoAAAa15097
echo cl.logfile = \'/marian/data2/mds/draco04Aug92.log\' >> /marian/data2/mds/D7
echo cl.keeplog = yes >> /marian/data2/mds/dracoAAAa15097
# initialize implementation STSDAS-CR-REMOVAL
echo stsdas >> /marian/data2/mds/dracoAAAa15097
echo wfpc >> /marian/data2/mds/dracoAAAa15097
echo combine.logfile = \'/marian/data2/mds/draco04Aug92.log\' >> /marian/data2/7

# invoke implementation STSDAS-CR-REMOVAL
echo combine @/marian/data2/mds/dracoAAAa15095 /marian/data2/mds/dracoAAAa150997
echo combine @/marian/data2/mds/dracoAAAa15093 /marian/data2/mds/dracoAAAa151017
# invoke package IRAF
echo logout >> /marian/data2/mds/dracoAAAa15097
/usr/stsci/irafx/unix/hlib/cl.csh < /marian/data2/mds/dracoAAAa15097 > /dev/null
rm /marian/data2/mds/dracoAAAa15097
echo created files /marian/data2/mds/draco04Aug92.log /marian/data2/mds/dracoAA1
```

SIGMA: SCIENTISTS' INTELLIGENT GRAPHICAL MODELING ASSISTANT

Dr. Richard Keller
Ames Research Center

August 11, 1992

Σ SIGMA: Scientists' Intelligent Graphical Modeling Assistant

Project Team

Richard Keller, David Thompson, Michal Rimon, Aseem Das
Artificial Intelligence Research Branch

Michael Sims, Yaron Gold
Information Sciences

Jennifer Dungan
Ecosystem Sciences

Chris McKay, Caitlin Griffith
Planetary Sciences

NASA *Ames Research Center*
Project contact: R. Keller, (415) 604-3388, keller@ptolemy.arc.nasa.gov

Outline

I. Introduction
- Motivation & goals
- Project plan
- Scope
- Approach

II. Example Model-building session
- Titan planetary atmospheric model
- Model-building interaction

III. Summary and future directions

Motivation

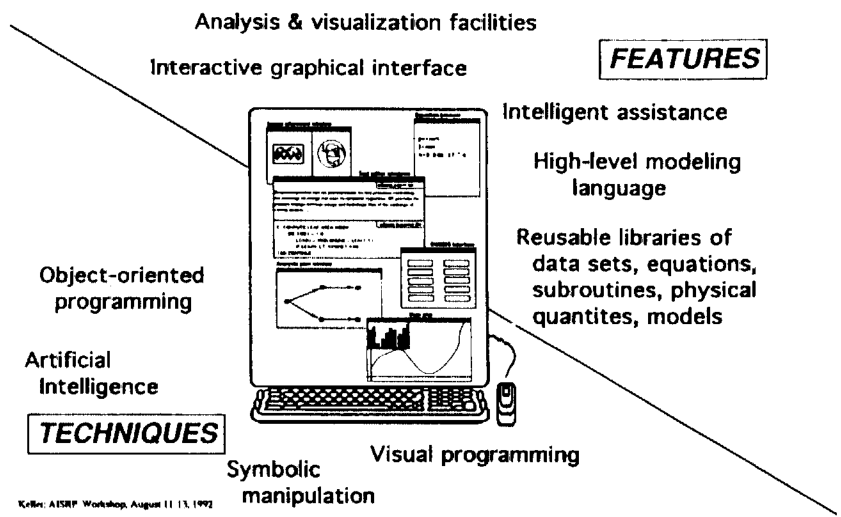
Model-building is essential for scientific advancement, yet time-consuming and error-prone

Problems
- Labor intensive to produce models
- Difficult to modify models
- Difficult to share and reuse models

Causes
- Obtuse, low-level code
- Wrong level of abstraction
- Important assumptions buried

2

Project Goal: Build a *specialized* software tool to assist in Scientific Model-building



Overall Project Plan

- Develop and implement a useful modeling tool for specific targeted applications
- Deploy tool to support scientists in Cassini and EOS missions
- Produce a generalized Scientific Modeling "shell" to handle a class of modeling tasks

What types of Models?

Set of algebraic & ordinary differential equations

- "Basic science" causal process models (vs. empirical correlational models)
- "Clean", simplified, micro-science models (vs. Complex, macro-science models)
- Data-poor environment (vs. data-rich environment)
- Deterministic (vs. uncertain / statistical)
- Quantitative (vs. qualitative)
- Spatial and temporal extent

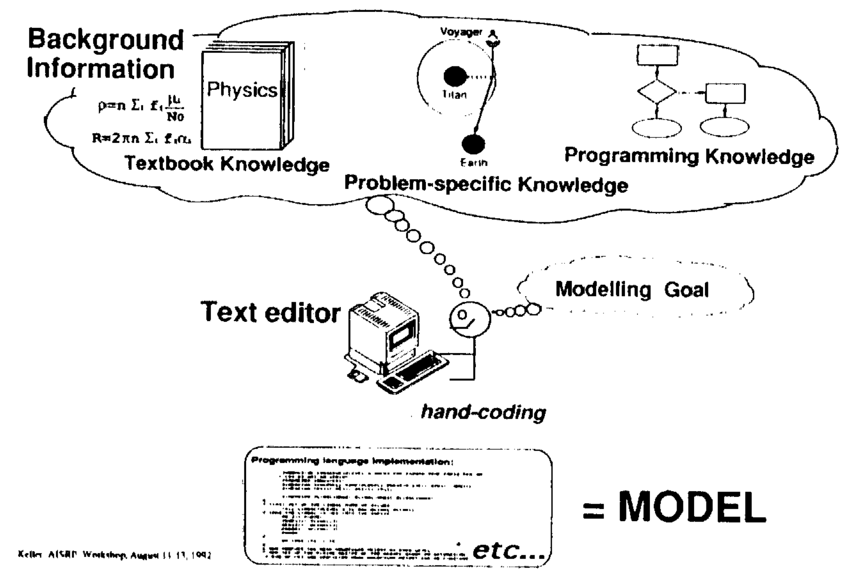
CURRENT MODELS

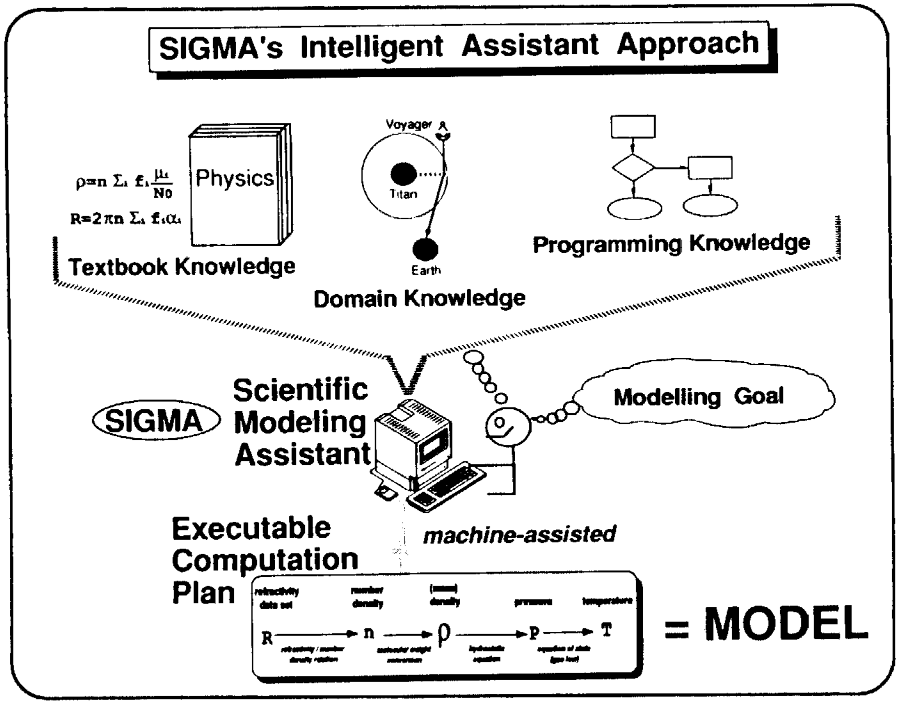
- TGM (Titan Greenhouse Model) C.P. Mckay- NASA Ames
- Forest-BGC (Ecosystem model) S.W. Running - U of Montana

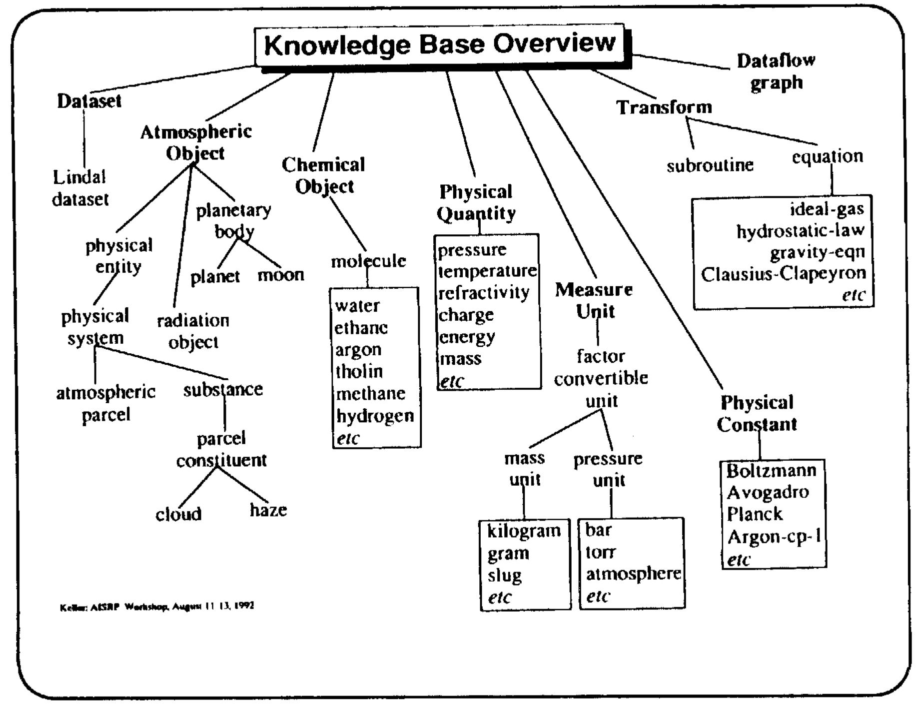
PROSPECTIVE MODELS

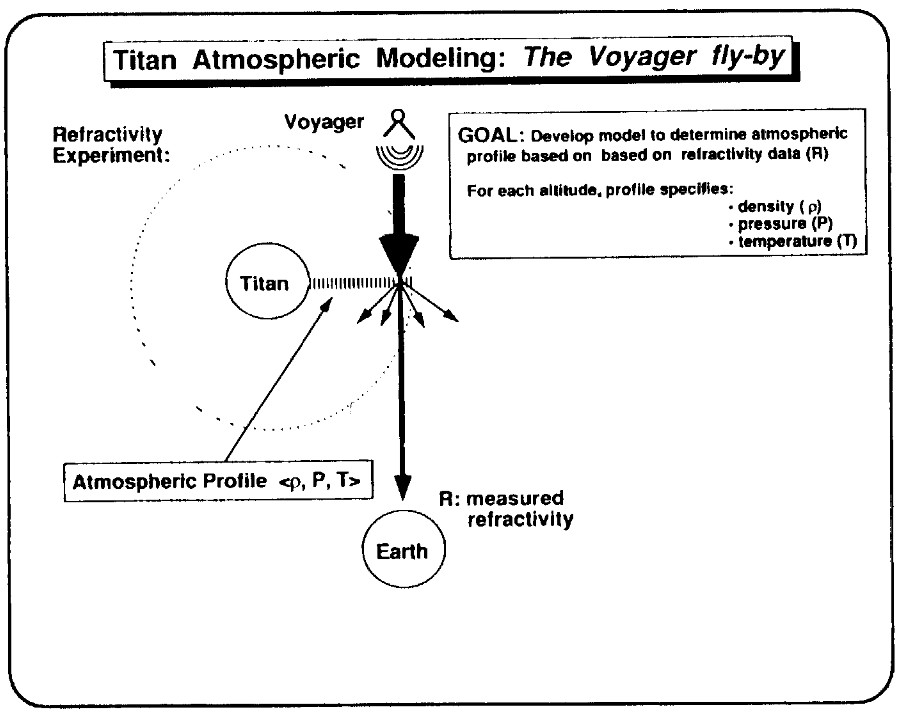
- Life Support System Models

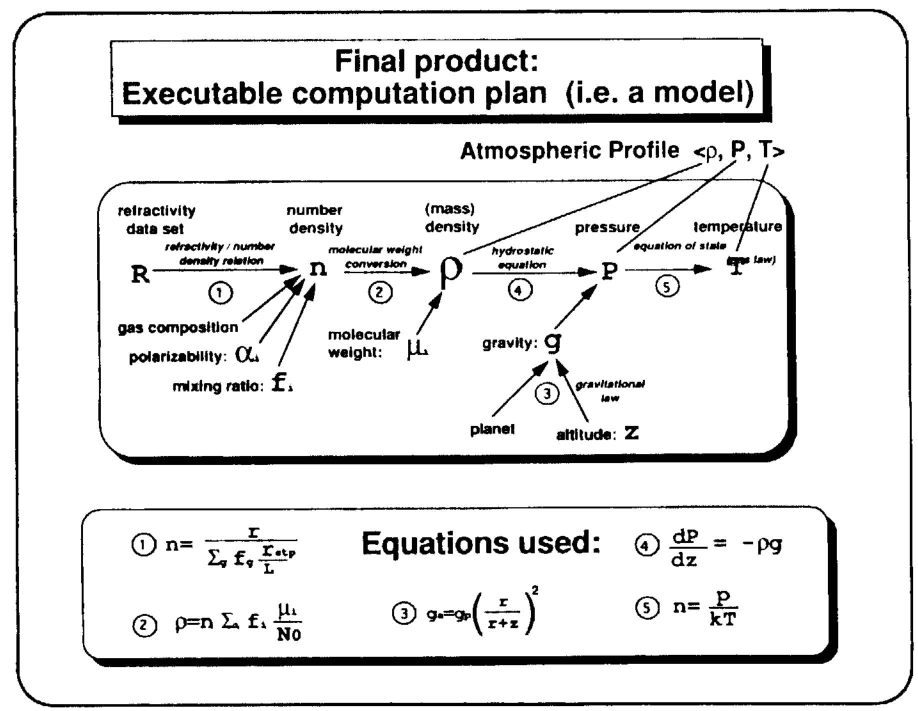
Manual Approach to Model-building




SIGMA's Intelligent Assistant Approach
ρ=n Σᵢ fᵢ μᵢ/N0
R=2πn Σᵢ fᵢαᵢ
Physics
Textbook Knowledge
Voyager
Titan
Earth
Domain Knowledge
Programming Knowledge
SIGMA
Scientific Modeling Assistant
Modelling Goal
Executable Computation Plan
machine-assisted
R → n → ρ → P → T
= MODEL


Knowledge Base Overview
Dataset
Lindal dataset
Atmospheric Object
physical entity
physical system
atmospheric parcel
planetary body
planet
moon
radiation object
substance
parcel constituent
cloud
haze
Chemical Object
molecule
water
ethane
argon
tholin
methane
hydrogen
etc
Physical Quantity
pressure
temperature
refractivity
charge
energy
mass
etc
Measure Unit
factor convertible unit
mass unit
kilogram
gram
slug
etc
pressure unit
bar
torr
atmosphere
etc
Physical Constant
Boltzmann
Avogadro
Planck
Argon-cp-1
etc
Transform
subroutine
equation
ideal-gas
hydrostatic-law
gravity-eqn
Clausius-Clapeyron
etc
Dataflow graph
Keller: AISRP Workshop, August 11 13, 1992


Titan Atmospheric Modeling: The Voyager fly-by
Refractivity Experiment:
Voyager
GOAL: Develop model to determine atmospheric profile based on based on refractivity data (R)
For each altitude, profile specifies:
• density (ρ)
• pressure (P)
• temperature (T)
Titan
Atmospheric Profile <ρ, P, T>
R: measured refractivity
Earth


Final product:
Executable computation plan (i.e. a model)
Atmospheric Profile <ρ, P, T>
refractivity data set
number density
(mass) density
pressure
temperature
refractivity / number density relation
molecular weight conversion
hydrostatic equation
equation of state (gas law)
R → n → ρ → P → T
gas composition
polarizability: αᵢ
mixing ratio: fᵢ
molecular weight: μᵢ
gravity: g
gravitational law
planet
altitude: z
Equations used:
① n = r / (Σ_q f_q r_stp / L)
② ρ = n Σᵢ fᵢ μᵢ / N0
③ g_z = g_p (r / (r+z))²
④ dP/dz = −ρg
⑤ n = P / kT

Model-building Steps

1. Modeling Scenario Setup

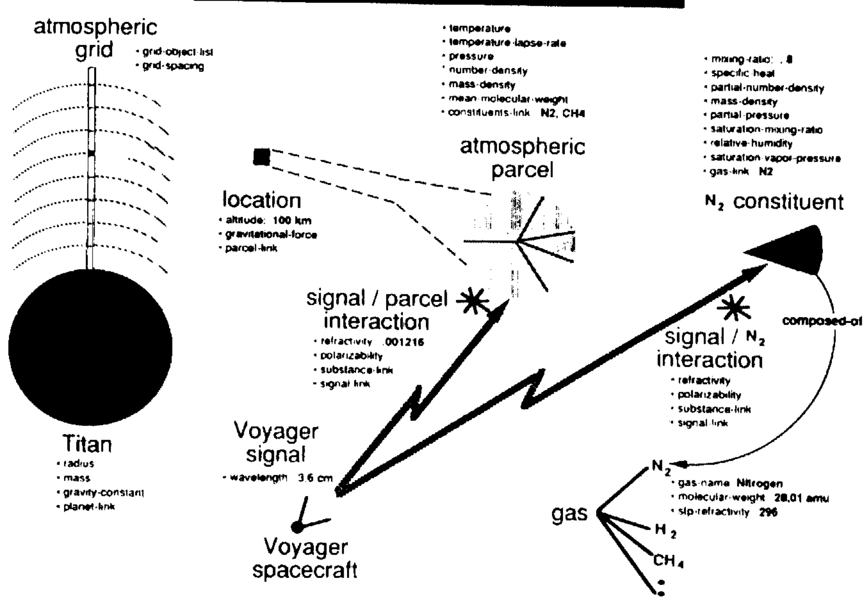
supported by interface {
2. Computation Plan Construction
3. Plan Execution

1. Modeling Scenario Setup

Build object-oriented representation of modeling scenario

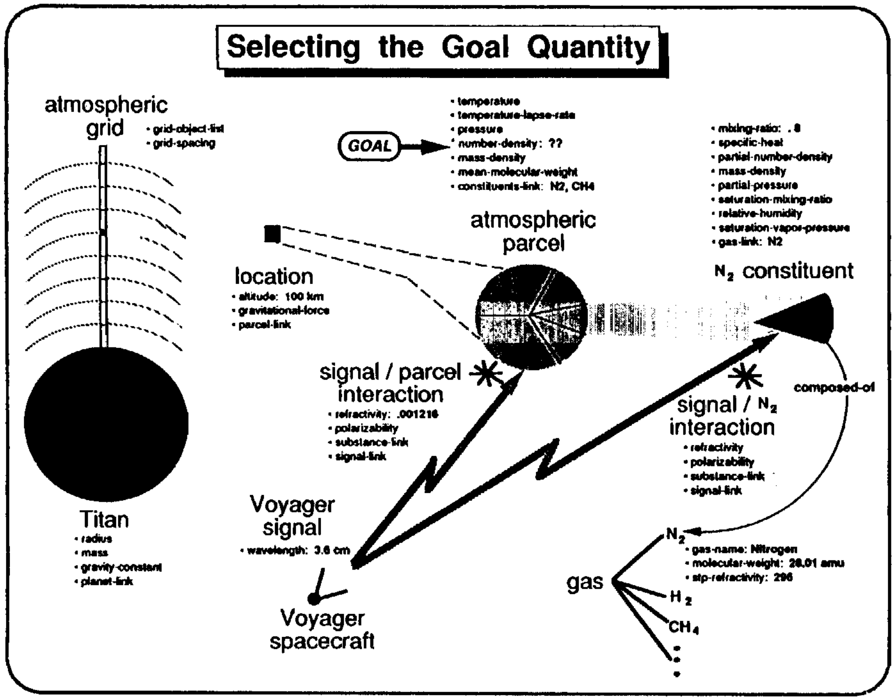
- Select the domain objects to be modelled
- Establish relationships between objects
- Initialize any known quantities associated with domain objects

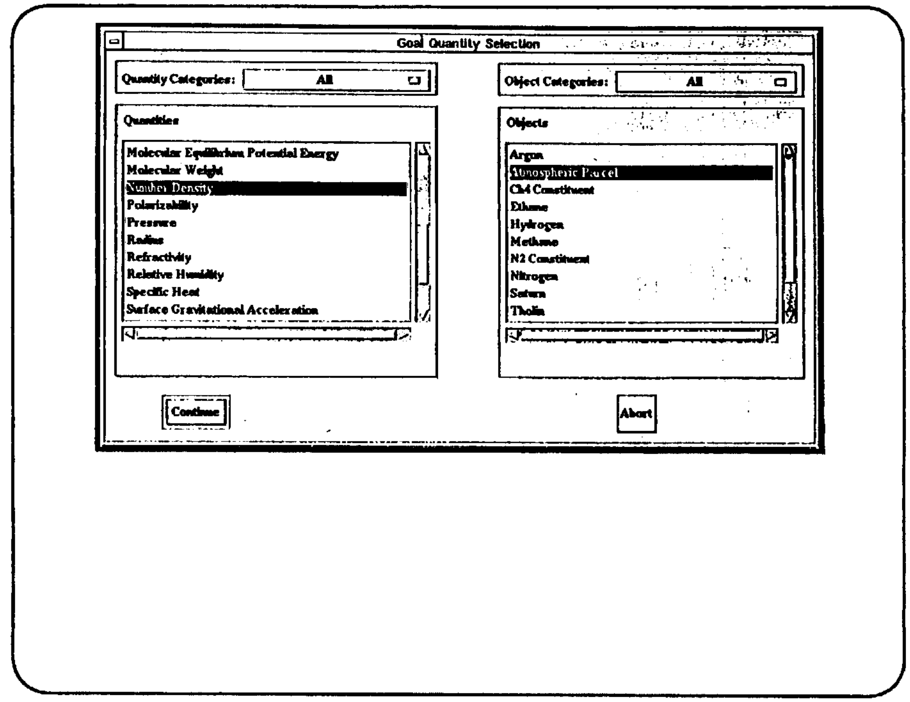
Initial Titan Modeling Scenario

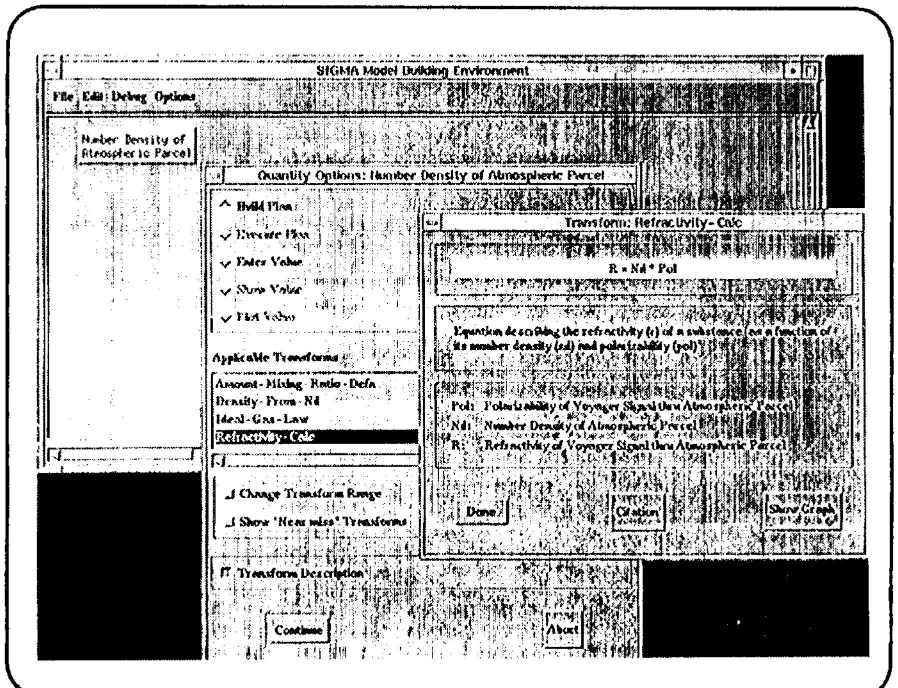


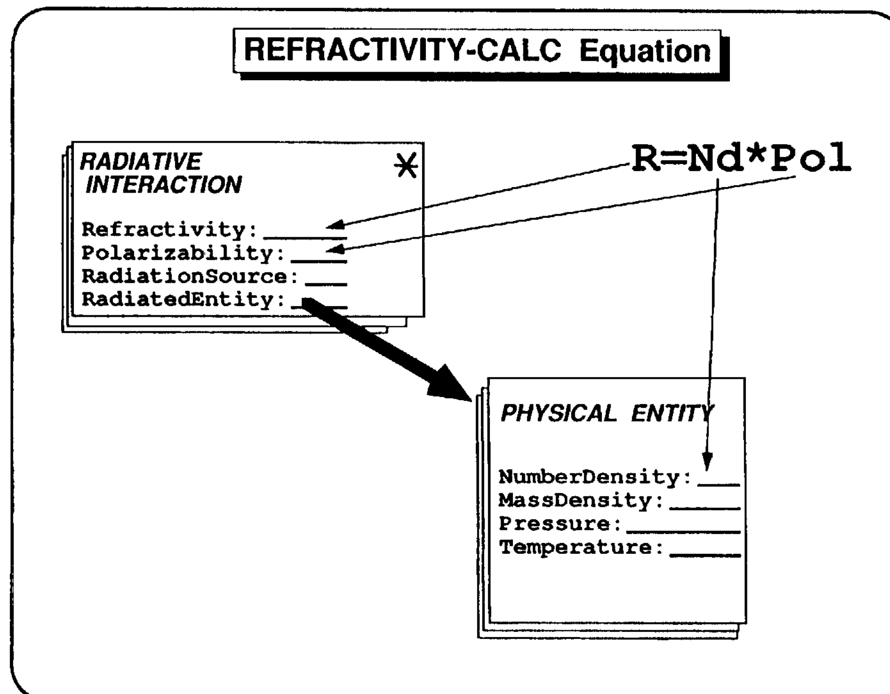
2. Computation Plan Construction

- Select a "goal quantity"
- Interactively devise a sequence of transforms to calculate the goal quantity

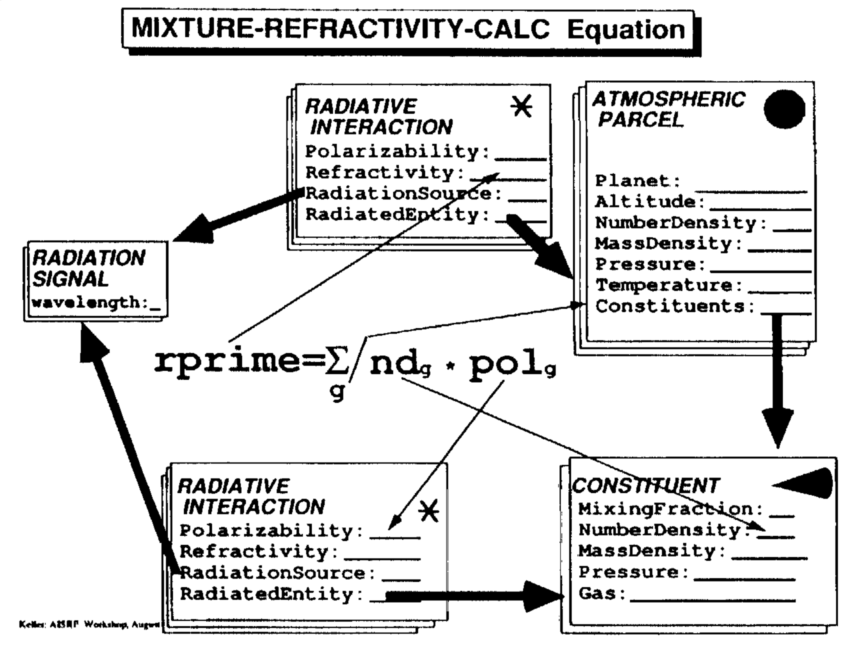


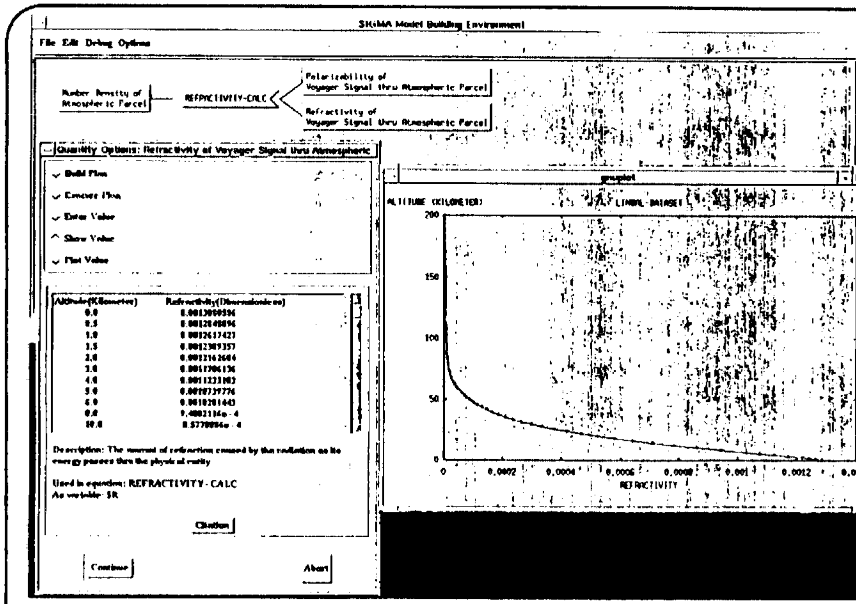


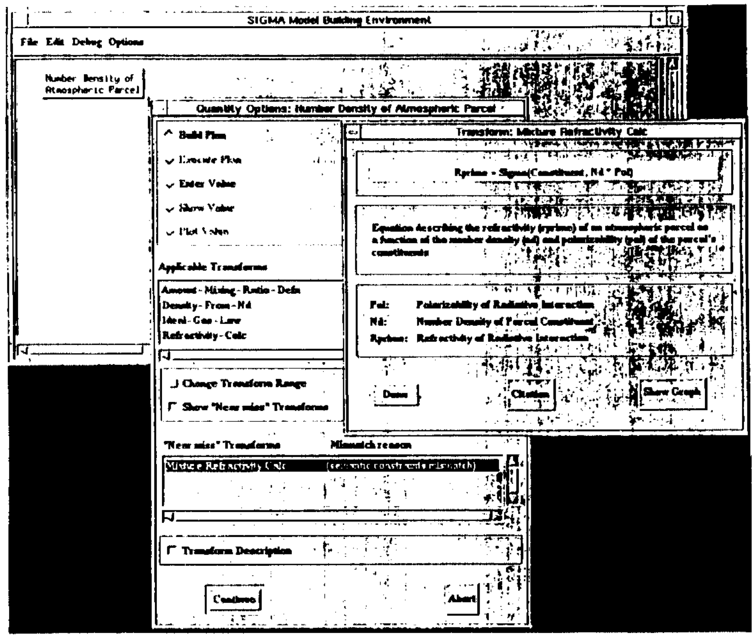


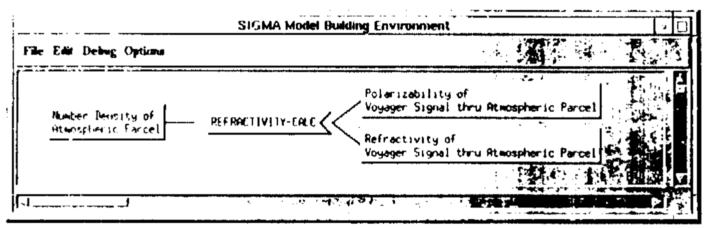


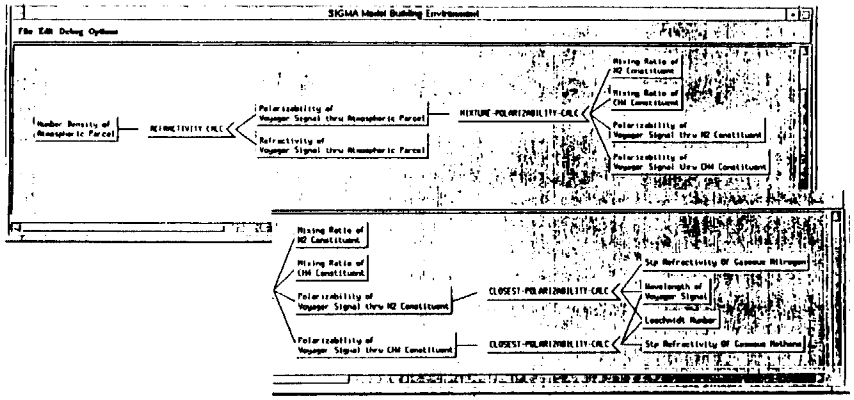
ORIGINAL PAGE IS
OF POOR QUALITY

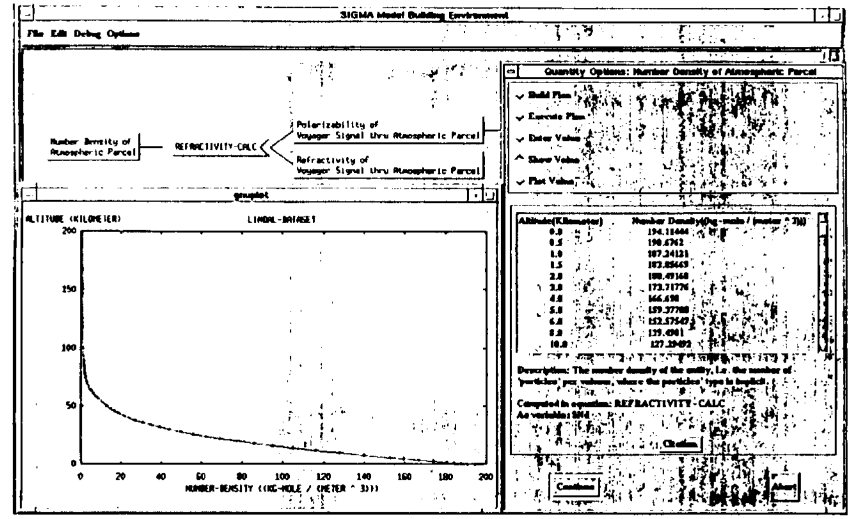

MIXTURE-REFRACTIVITY-CALC Equation
RADIATIVE INTERACTION
Polarizability:
Refractivity:
RadiationSource:
RadiatedEntity:
ATMOSPHERIC PARCEL
Planet:
Altitude:
NumberDensity:
MassDensity:
Pressure:
Temperature:
Constituents:
RADIATION SIGNAL
wavelength:
rprime=Σ_g nd_g * pol_g
RADIATIVE INTERACTION
Polarizability:
Refractivity:
RadiationSource:
RadiatedEntity:
CONSTITUENT
MixingFraction:
NumberDensity:
MassDensity:
Pressure:
Gas:
Keller: AISRP Workshop, August


SIGMA Model Building Environment
File Edit Debug Options
Number Density of Atmospheric Parcel
REFRACTIVITY-CALC
Polarizability of Voyager Signal thru Atmospheric Parcel
Refractivity of Voyager Signal thru Atmospheric Parcel
Quantity Options: Refractivity of Voyager Signal thru Atmospheric
Build Plan
Execute Plan
Enter Value
Show Value
Plot Value
Altitude(Kilometer)
Refractivity(Dimensionless)
Description: The amount of refraction caused by the radiation as its energy passes thru the physical entity
Used in equation: REFRACTIVITY-CALC
As variable: $R
Citation
Continue
Abort
gnuplot
ALTITUDE (KILOMETER)
REFRACTIVITY

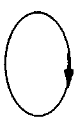

SIGMA Model Building Environment
File Edit Debug Options
Number Density of Atmospheric Parcel
Quantity Options: Number Density of Atmospheric Parcel
Build Plan
Execute Plan
Enter Value
Show Value
Plot Value
Applicable Transforms
Amount-Mixing-Ratio-Defn
Density-From-Nd
Ideal-Gas-Law
Refractivity-Calc
Change Transform Range
Show "Near miss" Transforms
"Near miss" Transforms
Mismatch reason
Mixture Refractivity Calc
Transform Description
Continue
Abort
Transform: Mixture Refractivity Calc
Rprime = Sigma(Constituent, Nd * Pol)
Pol: Polarizability of Radiative Interaction
Nd: Number Density of Parcel Constituent
Rprime: Refractivity of Radiative Interaction
Done
Citation
Show Graph


SIGMA Model Building Environment
File Edit Debug Options
Number Density of Atmospheric Parcel
REFRACTIVITY-CALC
Polarizability of Voyager Signal thru Atmospheric Parcel
Refractivity of Voyager Signal thru Atmospheric Parcel





3. Computation Plan Execution



- Determine whether any tranforms are fireable
- Fire those transforms

(A transform is fireable if all of its inputs are known)

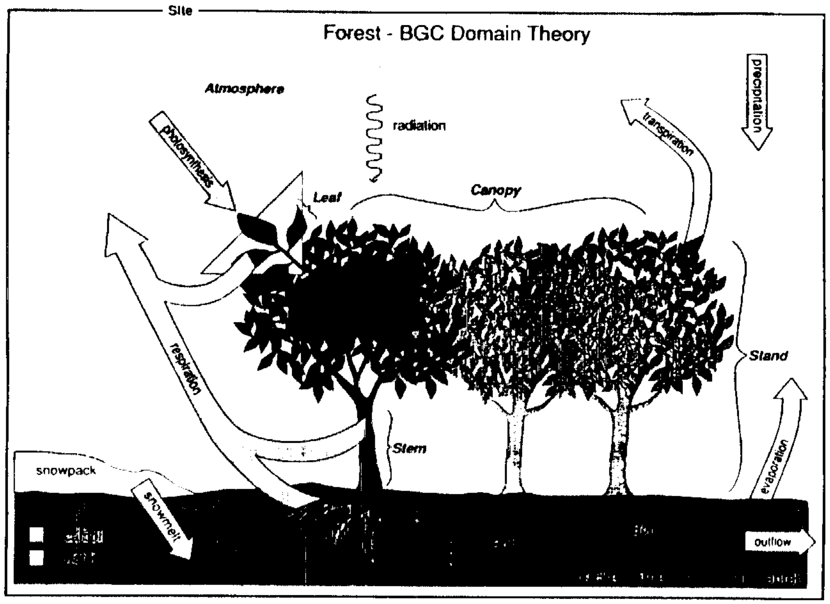
<u>To fire a transform</u>:

- Convert input values to a common scientific unit
- Execute transform
- Store output value

SIGMA Benefits for the Modeller

- Immediate feedback
- Visualization of model and data
- Sensitivity analysis support
- Intelligent support for equation selection
- Automatic scientific units maintenance
- <u>Reuse</u> of pre-existing
  - *transforms* (equations and subroutines)
  - *objects and attributes*
  - *models*
- Automatic support for handling data structure

Initial Forest-BGC Modelling Scenario



FY93 Milestones

- Finish Titan temperature determination model
- Implement Forest-BGC modeling scenario
- Build graphical equation-entering facility
- Deliver experimental prototype to science users for testing and evaluation

OAST AI PROGRAM

Dr. David Thompson
Ames Research Center

August 11, 1992

OAST AI Program

- $13M / YR PROGRAM 7 NASA CENTERS
- EARTH and SPACE SCIENCE DOMAINS are a MAJOR FOCUS of the PROGRAM at NASA AMES and JPL

Intelligent Scientific Laboratory Instruments

Bayesian Data Analysis -- AutoClass

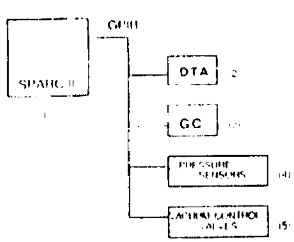
PI-in-a-Box

Multi-Agent Planning for Heterogeneous Registration

Reactive Planning, Scheduling, and Control

Spacecraft Health Automated Reasoning Prototype

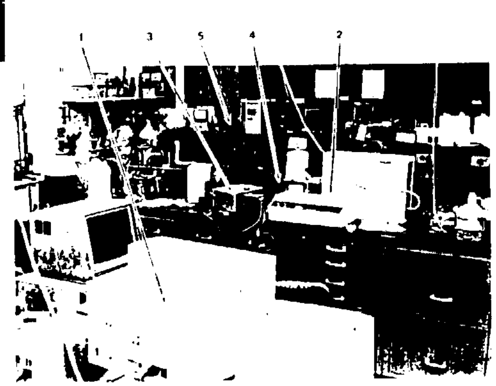
Scientific Analysis Assistant

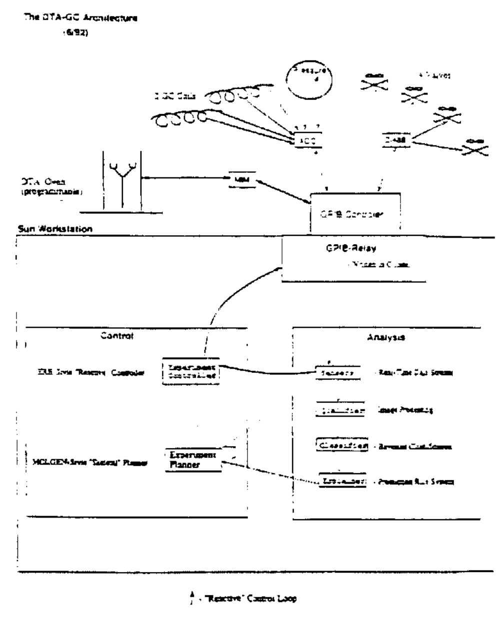


DIFFERENTIAL THERMAL ANALYZER GAS CHROMATOGRAPH

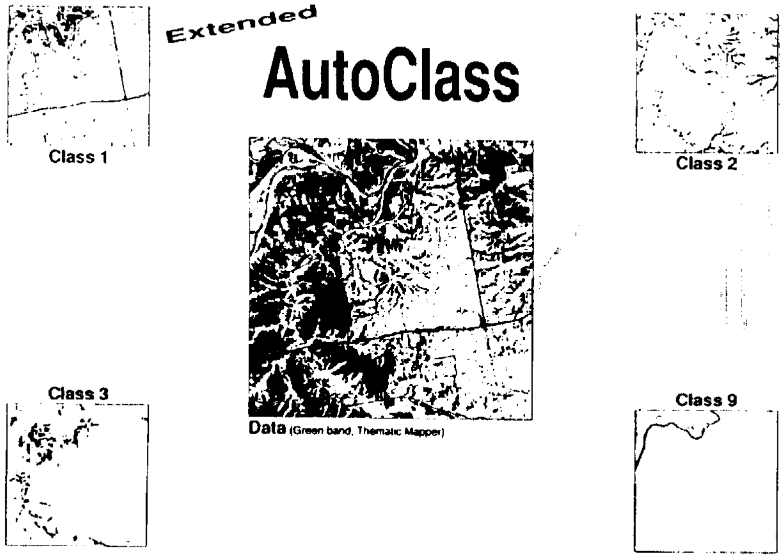
AUTONOMOUS CONTROLLER AND ANALYSIS SYSTEM

- SYSTEM IDENTIFIES UNKNOWN SAMPLES AND MODIFIES EXPERIMENT STRATEGIES





Extended AutoClass



The extended AutoClass implementation was applied to data from Landsat's Thematic Mapper data from a study area (F.I.F.E.) in Kansas, shown in center panel. For each pixel, information from 7 spectral bands was used to build the class descriptions, including the correlations between the bands. Only 4% of the pixels were used to build the classification because of time limitations. Examples of some of the classes found are shown in the surrounding panels, where each pixel (not just the 4% sample) has been assigned to the class in which its membership is maximum. The spatial location of the pixels was not used to inform the classification, but note that it shows up clearly in the classes

Data Driven Earth and Atmopheric Modelling Using Bayesian Reasoning

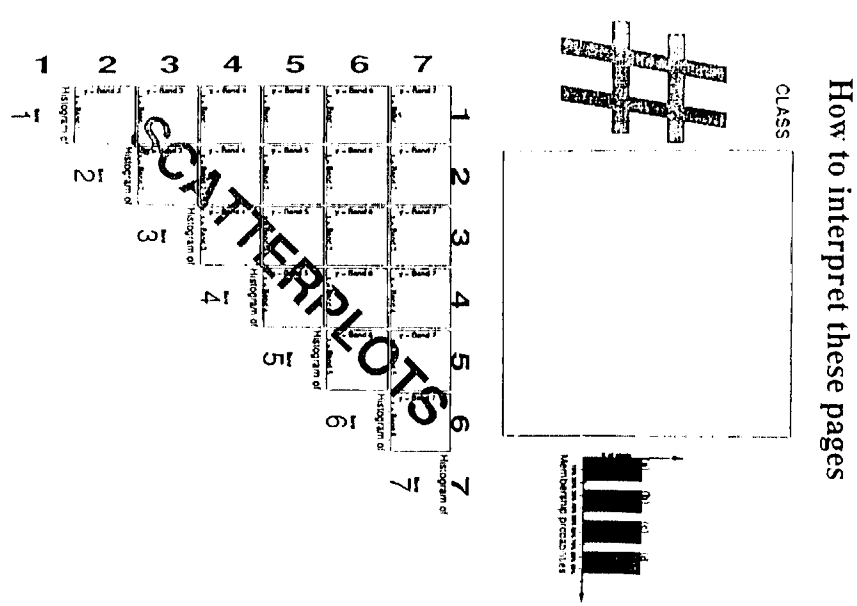
— Peter Cheeseman, Bob Kanefsky, John Stutz, Robin Hanson, Will Taylor

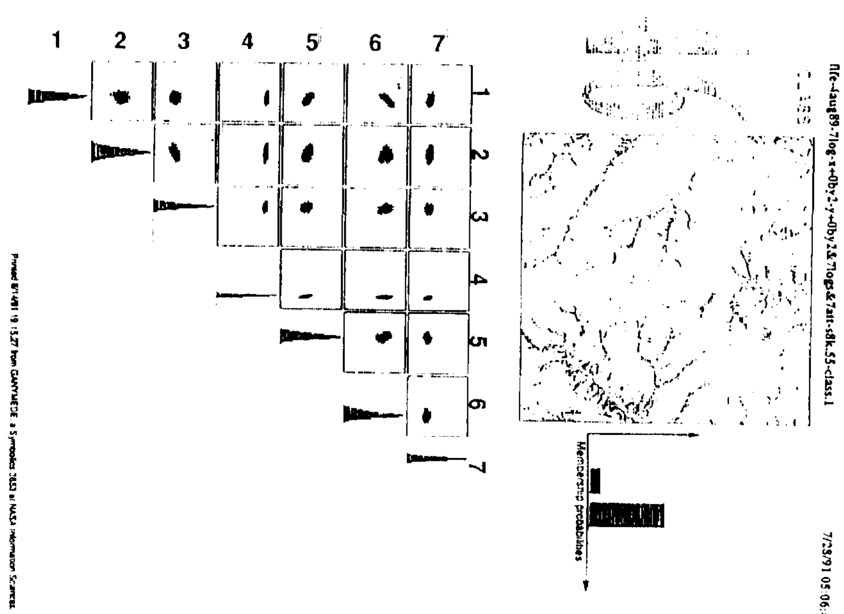
This project is focused on using Bayesian probability theory to extract the most probable information inherent in data.

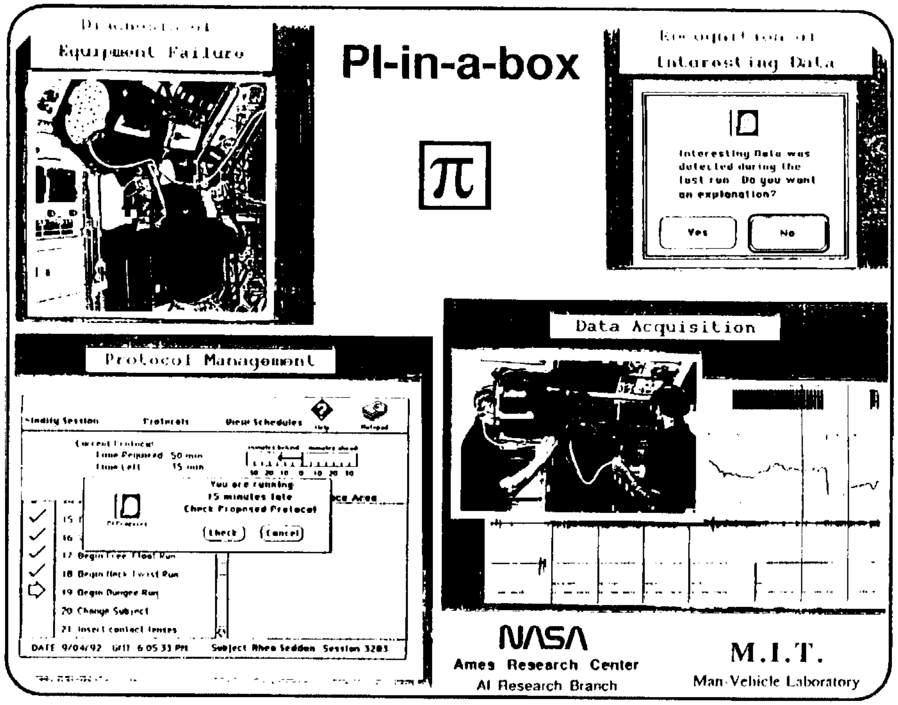
Applications of this work include:

- Finding the most probable clusters or classes of ground features, based on satellite data.
- Automatically registering multiple images of the same ground area (taken by the same instrument or some highly correlatable instrument) and using multiple images to form *higher* resolution data than available in each individual data set.
- Automatically discovering new classes of observations (e.g., detecting anomalies).
- Finding the most probable correlations e.g., does accumulation of certain gases really correlate to certain atmospheric phenomenon?

How to interpret these pages







Motivation

Time and resource constraints severely limit flexibility during space experimentation:

- PI is physically distant from experiment
- Communication bandwidth availability is often insufficient or not timely enough
- Quantity and complexity of experiments impact productivity
- Crew availability and training overhead are major issues in the Space Station environment

ASA Overview

- **Objectives:**

  To improve the scientific return of experiments performed in space

  To evaluate Data Management accomodations needed to support payloads

- **Approach:**

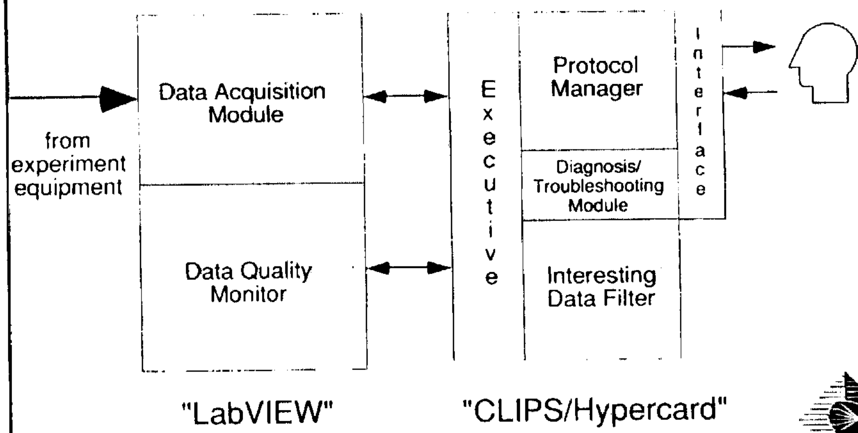
  Use expert systems technology to encode domain and experiment knowledge of the Principal Investigator and make it available to payload and mission specialists

  Operationally evaluate the system in Spacelab domain and in Advanced Architecture Testbed

Functions of the ASA

- Capture, reduce, and archive experimental data
- Monitor data quality and help diagnose problems with equipment when experimental data is erratic or poor
- Identify and permit investigation of "interesting" data
- Suggest protocol changes that would result in better utilization of remaining time

System Architecture

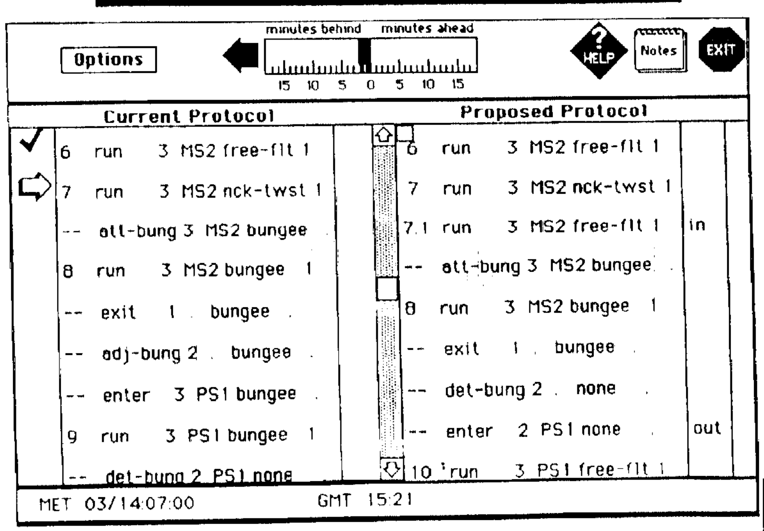


Hypothetical ASA Scenario

- The dome experiment, with two subjects, is running slightly behind schedule.
- Subject 1 had exhibited "interesting data" on the previous day.
- Subject 2 had exhibited erratic data during the previous session that same day.

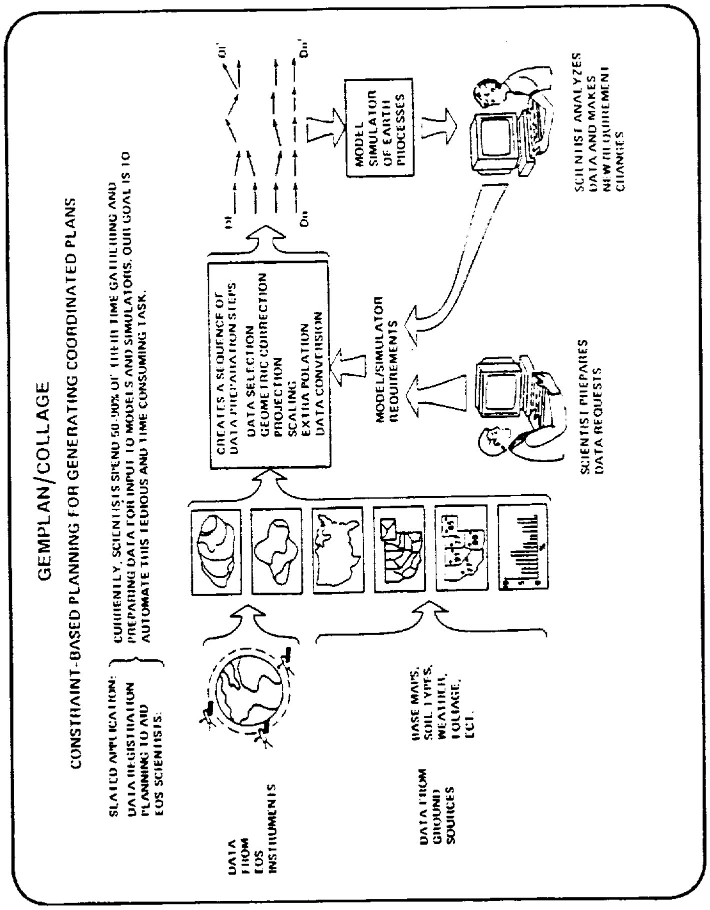
How should the protocol be refined to maximize the scientific return of this session?

The "Proposed" Protocol



Support of SLS-1 Mission

- **Pre-flight baseline data collection:**
  - system used to collect and analyze data from Rotating Dome experiment in the Baseline Data Collection Facility at JSC on L-150, L-75, L-45, L-30, and L-15 sessions
- **Ground support during flight experiment:**
  - system used in the Science Monitoring Area at JSC to collect and analyze in-flight data from the Dome experiment downlinked from Spacelab
- **Post-flight data collection:**
  - system used at Dryden to collect and analyze data from the Dome experiment on R+0, R+1, R+2, R+4, R+7, and R+10 sessions



REGISTRATION PLANNING FOR HETEROGENEOUS DATA SETS

Amy Lansky, Andrew Philpot, Jennifer Dungan

The primary concern of many earth scientists is to build and verify models of earth ecosystems. However much time is spent on tedious task of *selecting and preparing data* for input into these models.

GIVEN:

N desired data set descriptions for specific locations and and time periods

CONSTRUCT:

A data-access and transformation plan that yields data sets that:

- sufficiently meets time/space requirements
- supply data of the right type
- are all within the same coordinate system (pixel/pixel line up):
  - same scale
  - same projection
  - same or adequate density of information

Accomplishments

The system worked under realistic conditions.

- Collection and archival of downlinked data
- Quick-look analysis and summary of data
- Generation of potential new protocols

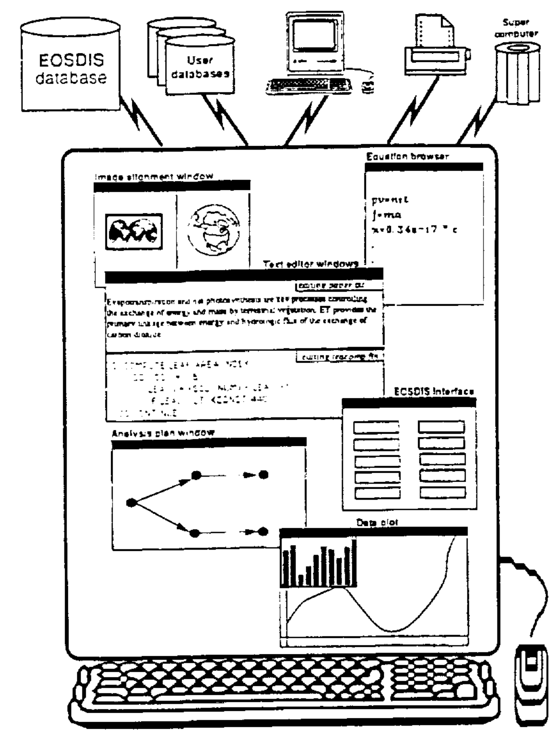
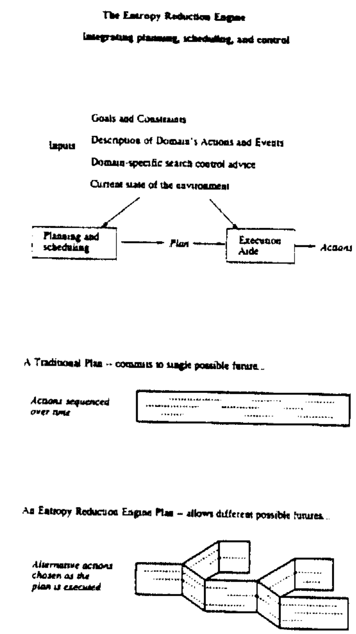


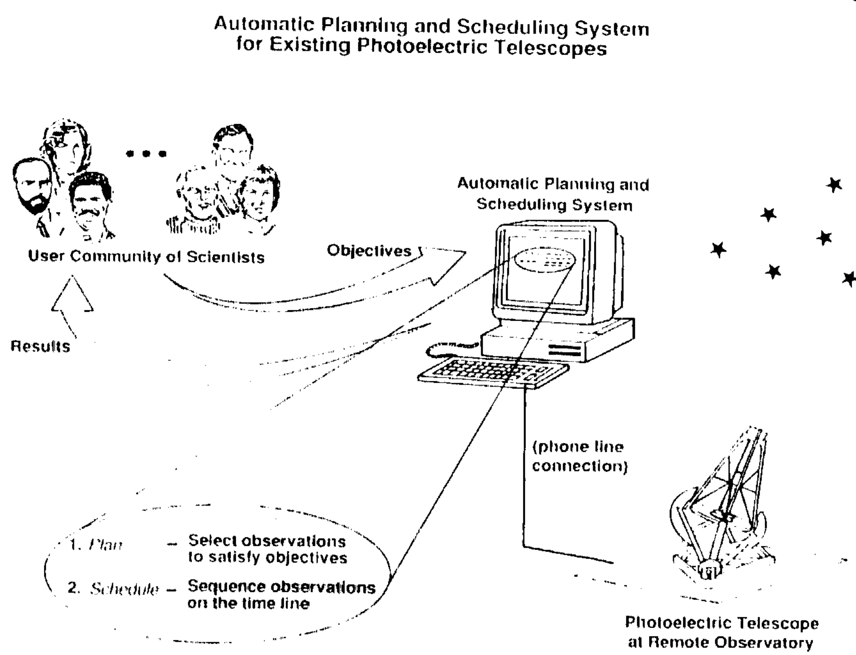
Figure 1. Data Analysis Workbench

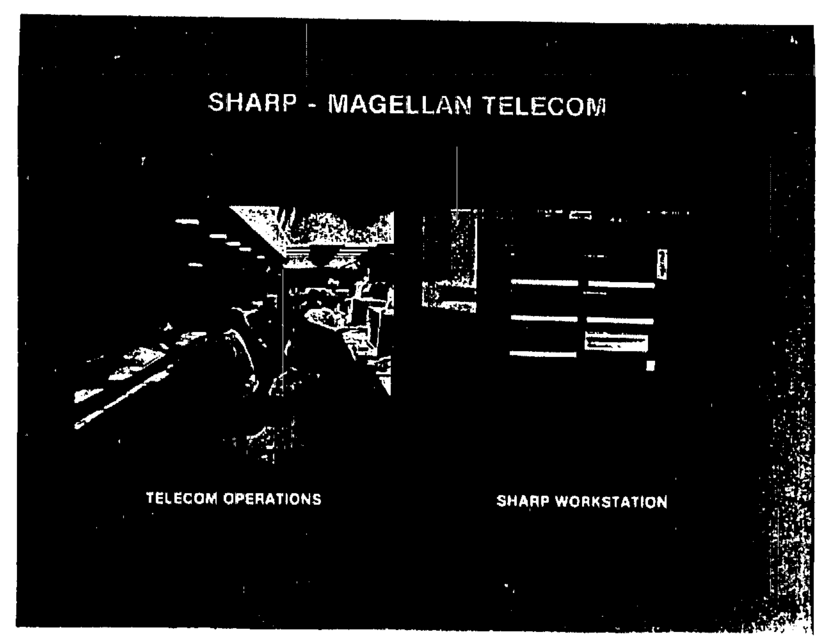
ORIGINAL PAGE IS OF POOR QUALITY

BENEFITS OF AUTOMATION:

- Allows more complete and correct exploration of registration possibilities.
- Automates a busy-work task.
- Speeds up analysis, allowing for more science "quality time."
- Provides an automated interface to EOS-DIS and models.
- Provides a framework for diagnosing errors or distortions in the data.
- Allows for reuse of registration plans.

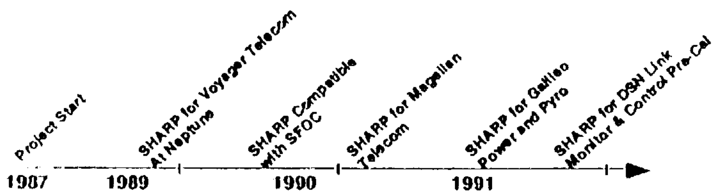






SHARP PROGRESS UPDATE

JPL



Evaluation Prototype	Reusable Kernel	Pilot Installation	Pilot Installation	Underway
Shallow Diagnosis	Constraint Based Diagnosis	Deeper Telecom Diagnosis	Embedded Diagnosis Shell	
30 Sec. to Diagnose	1.5 Sec. to Diagnose	"Anytime" Diagnosis	Effectively Real-time	
Max ~ 100 RT Channels	Max ~ 10K RT Channels	Capacity to Spare	Capacity to Spare	
LISP Machine	Sun and SFOC Compatible	Installed in Magellan Ops	Installed in Galileo Ops	

SAA -- SkICAT

Application: **Cataloguing sky objects from the 2nd. Palomar Sky Survey (POSS-II).**

System: **SkICAT (Sky Image Classification and Archiving Tool)**

Goal: take as input a digitized photographic plate produce as output catalog entries for all objects in image.

Objects:
star (s)
star with fuzz (sf)
galaxy (g)
artifact (long)

JPL

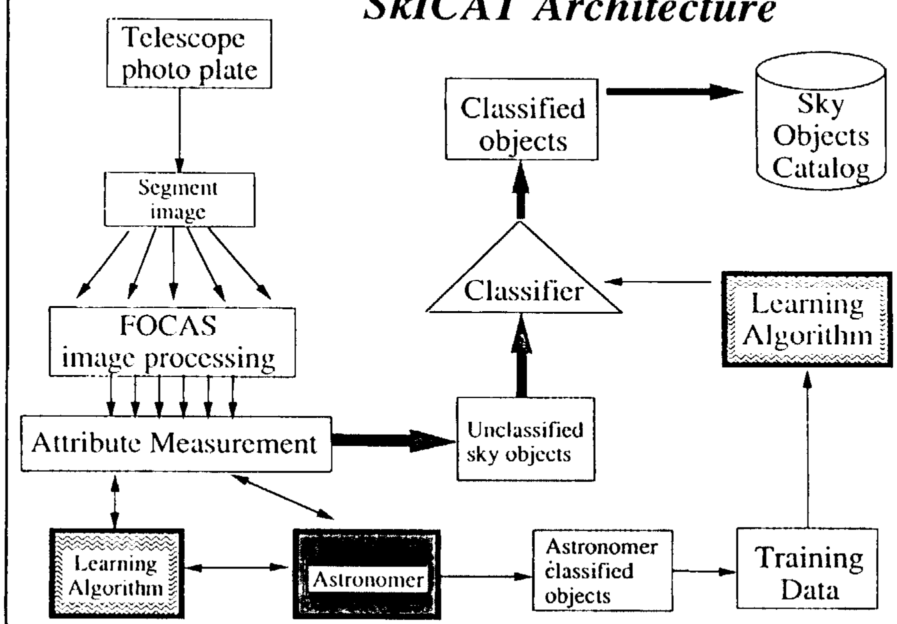
SAA -- SkICAT

Problems:

- Too voluminous for human processing (≈ 1800 plates, each having 23,000x23,000 pixel, ==> 9×10^{11} bytes of data
- Most objects too faint for recognition by inspection, need automated methods.
- Cataloguing is tedious, expensive step needed before further scientific analysis can be performed.

JPL

SkICAT Architecture

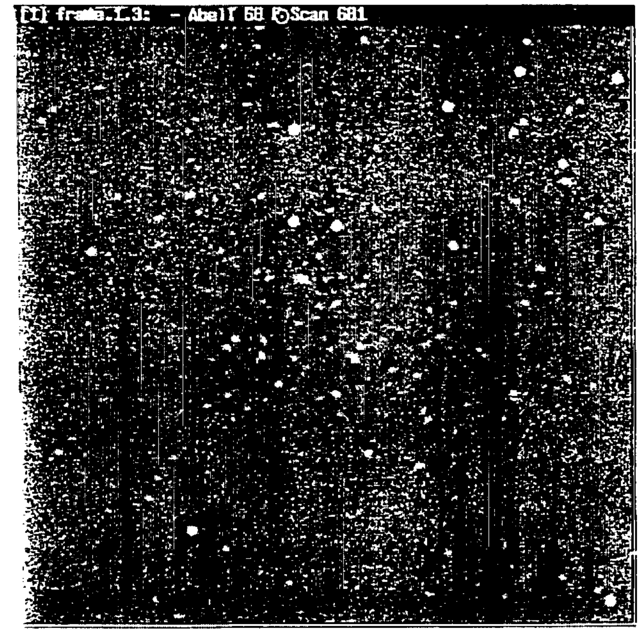


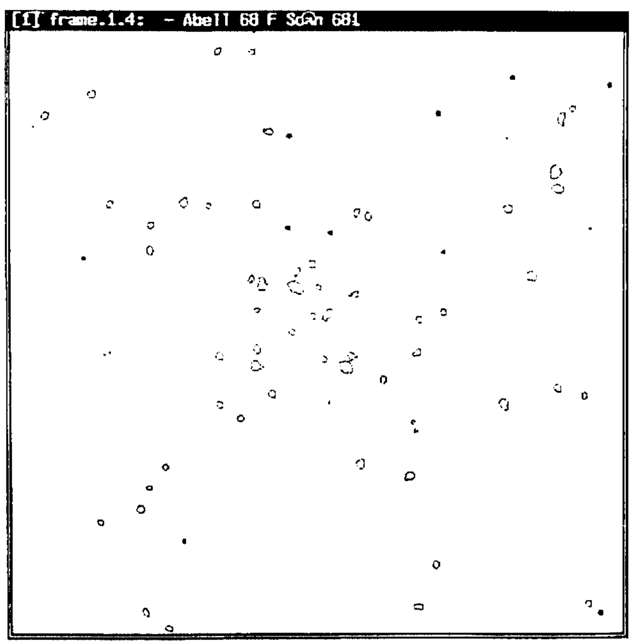
SAA -- SkICAT

Benefits:

- Speed up catalog generation by one to two orders of magnitude (unrealistic to perform manually).
- Reduce cost of cataloguing survey images by equivalent of tens of astronomer man-years.
- Classify objects that are at least one magnitude fainter than catalogs to-date.
- Generate catalogs with much richer content.
- Towards an objective, reliable automated classification method...

JPL





PARALLEL ALGORITHMS FOR DATA COMPRESSION

Dr. James A. Storer
Brandeis University

August 11, 1992

Parallel Algorithms for Data Compression

James A. Storer
Computer Science Dept.
Brandeis University

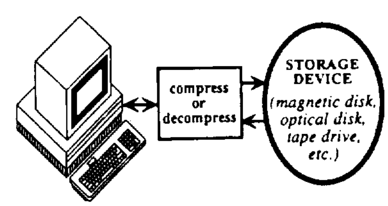
Joint Work With:

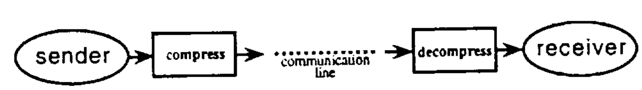
- *B. Carpentieri* (Ph D Student)
- *M. Cohn* (Faculty)
- *C. Contantinescu* (Ph D Student)
- *S. De Agostino* (Ph D Student)
- *E. Lin* (Post-Doc)
- *Q. Ye* (Ph D Student)
- *R. Zito-Wolf* (Ph D Student)

Outline

- Introduction
- Systolic Algorithms for Lossless Compression
- Lossless Compression Hardware
- NC (Poly-Log) Algorithms for Lossless Compression
- Vector Quantization for Image Compression
- On-Line Adaptive Vector Quantization
- Image Compression Visualization Tools and Experiments
- Image Compression Hardware
- Video Displacement Estimation
- Real-Time Video Compression Hardware
- Current Research

Applications of Data Compression





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.)

Lossless Data Compression with On-Line Dynamic Textual Substitution

Idea:

Maintain a dictionary of strings that have occurred frequently in the past and replace new occurrences of these strings by their index in the dictionary.

In a "lock-step" fashion, the encoder and decoder are constantly changing their dictionaries to best reflect the data seen thus far.

Data Compression with On-Line Dynamic Textual Substitution

(1) Initialize the *local dictionary* D to have one entry for each character of the input alphabet.

(2) repeat forever

- (a) {Get the current match string s:}
  - Use a *match heuristic* MH to read s from the input.
  - Transmit $\lceil log_2|D|\rceil$ bits for the index of s.
- (b) {Update D:}
  - Add each of the strings specified by an *update heuristic* UH to D (if D is full, use a *deletion heuristic* DH to make space).

Generic Encoding Algorithm

(1) Initialize D by performing Step 1 of the encoding algorithm.

(2) repeat forever

- (a) {Get the current match string s:}
  - Receive $\lceil log_2|D|\rceil$ bits for the index of s.
  - Retrieve s from D and output the characters of s.
- (b) {Update D:}
  - Perform Step 2b of the encoding algorithm.

Generic Decoding Algorithm

The Match Heuristic

- *Greedy*: Read the longest match possible.
- *Lookahead*: Employ a lookahead buffer to check if taking a shorter match now will pay off with better compression later.
- *Special Characters*: Read the longest match that ends with a special character (which can be specified in advance or dynamically learned).
-
-
-

Sliding Window Data Compression

Idea: A form on-line textual substitution where the dictionary is just a window of the last n characters and instead of pointers being simple indices, they are (*displacement, length*) pairs that indicate a substring of the window. The update and deletion heuristics are to just "slide" the window. Any match heuristic that reads a string that is a substring of the window will do; usually the greedy heuristic is used (and is provably optimal when all pointer have the same size).

Notes:

- To insure that a match of at least one character can always be found, a pointer value is reserved for each character of the alphabet.
- In practice, it pays to use "fancier" methods of encoding pointers because the distribution of pointer values (particularly the length field) tends not to be uniform. Also, better methods of coding pointers avoid the inefficiency of having to divide the pointer into two fields.

Efficient Serial Implementation:

- If the window extends back to the beginning of the input string, a simple linear time implementation is to build a position tree as you go (using McCreight's Algorithm) and compute longest matches by walking down from the root to a leaf and then matching as much additional input as possible.
- When the window does not extend all the way to the beginning of the input string, the position tree data structure can be modified to allow deletion of strings, or three overlapping copies of the position tree can be employed.

Update and Deletion Heuristics for Dynamic Dictionaries

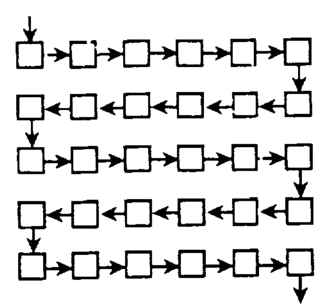
Update Heuristic (DH):

- *First Character* (FC): Add the last match concatenated with the first character of the current match.
- *Identity* (ID): Add the last match concatenated with the current match.
- *All Prefixes* (AP): Add the set of strings consisting of the last match concatenated with each of the prefixes of the current match.

Deletion Heuristic (DH):

- *Freeze when Full* (FREEZE): Once the dictionary becomes full it is "frozen" and remains the same from that point on.
- *Restart Periodically or when Compression Drops* (RESTART): Periodically or when compression drops, remove all elements of the dictionary (except the characters of the input alphabet).
- *Least Frequently Used* (LFU): Delete the string that has been matched least frequently.
- *Least Recently Used* (LRU): Delete the string that has been matched least recently.
- *Swap when Full* (SWAP): Keep two dictionaries. When the *primary* dictionary becomes full, start learning new entries in the *auxiliary* dictionary but continue compressing data with the primary dictionary. From this point on, each time the auxiliary dictionary becomes full, the roles of the primary and auxiliary dictionaries are reversed, and the secondary dictionary is reset to be empty.

Systolic Pipes

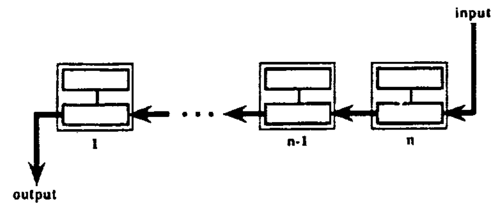


- All processors are identical and the length of connections between adjacent processors can be bounded by a constant.
- The structure can be laid out in linear area and power and ground can be routed without crossing wires.
- The layout strategy can be independent of the number of chips used. A larger pipe can be obtained by placing as many processors as possible on a chip and then, using the same layout strategy, placing as many chips as possible on a board.

Systolic Pipe for the

Static Dictionary Method

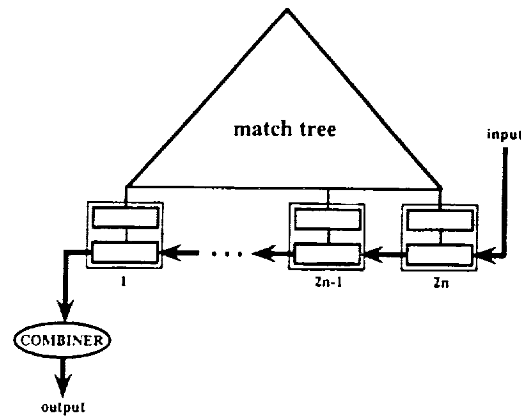
(the dictionary is fixed in advance)

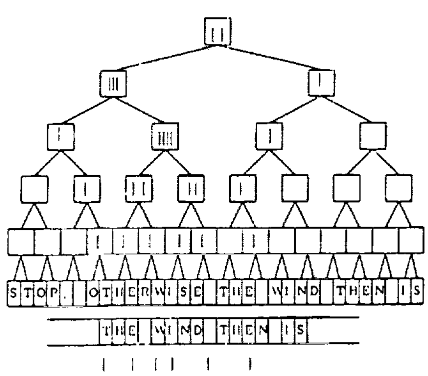


Systolic Architecture for the

Sliding Window Method

("match tree" architecture)



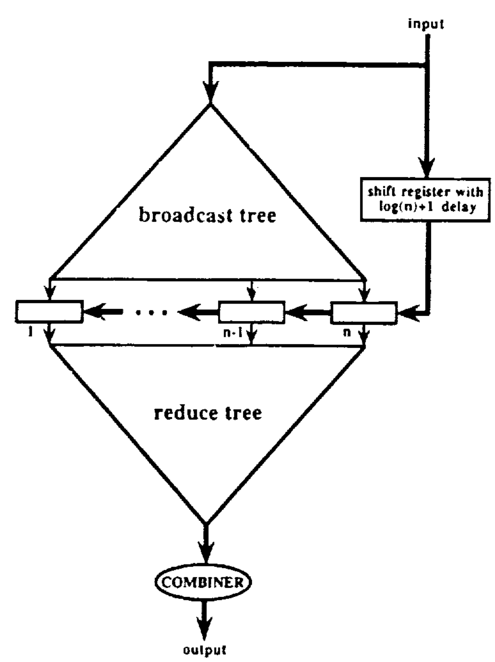


	INPUT	ACTION
1.	Nothing	None
2.	[	Send [to parent
3	]	Send] to parent
4	][	Send][to parent
5	[]	Compute length and send to processor that sent it
6.	[][	Handle first two as in 5 and send other [to parent
7	][]	Send first] to parent and handle last two as in 5
8.	][][	Send first and last to parent and handle as in 5.

Systolic Architecture for the

Sliding Window Method

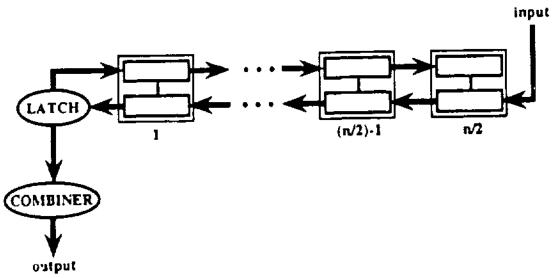
("broadcast-reduce" tree architecture)



Systolic Pipe for the

Sliding Window Method

("wrap" architecture)



Systolic Pipe for the

Dynamic Dictionary Method

(Uses a variant of the ID update heuristic that forms matches in a "bottom-up" fashion.)

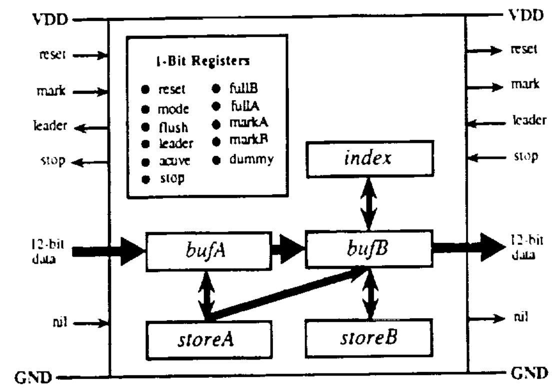
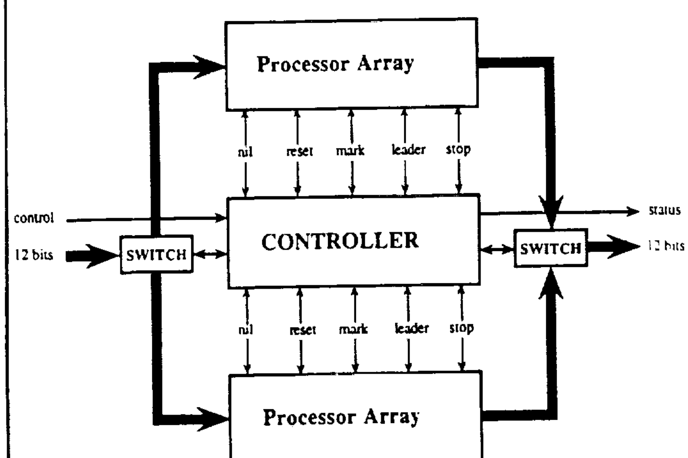


Diagram of a Single Cell

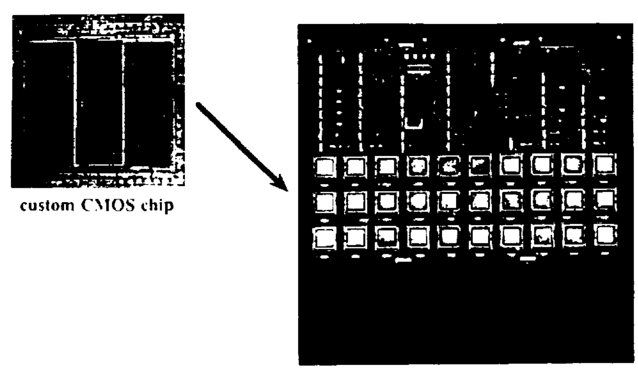
(4,096-256-1)=3,839 cells form a complete array)

Board Level Design that Implements the

Swap Deletion Heuristic



Real-Time Adaptive Lossless Compression Hardware



custom CMOS chip

compression / decompression board

VME Board:
- 20 mhz clock
- 30 custom chips (1.0u double-metal CMOS)
- 160 million bits per second throughput (one byte is processed on each clock cycle)

HiPPI Board:
- 100 mhz clock
- 15 custom chips (.8u double-metal CMOS)
- 800 million bits per second throughput

Sub-Linear Algorithms for Text Compression

Idea: Fast sub-linear algorithms that can be programmed on existing massively parallel machines.

Model of Input-Output: For a dictionary of size n, characters arrive in blocks of size n to the encoder and leave in blocks of size n from the decoder.

Note: Dynamic dictionary compression is P-Complete!

Static dictionary or sliding window: We present a simple algorithm for greedy or *optimal parsing* that can be implemented in $O(log(n))$ time with $O(n^2)$ processors.

(In fact, for any $0 < epsilon \leq \frac{1}{2}$, this algorithm can be implemented in $O(\frac{1}{\epsilon}log(n))$ time with $O(n^{1+\epsilon})$ processors.

Example:

input string= *a a b a a a b b b a a a b a a a*

dictionary= *a, b, aa, aab*

greedy parsing= *aab, aa, a, b, b, b, aa, a, b, aa, a*

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
4	3	4	6	8	7	8	9	10	12	14	13	14	16	17	17
4	3	4	6	8	7	8	9	10	12	14	13	14	16	17	17
6	4	6	7	9	8	9	10	12	13	16	14	16	17	0	0
4	3	4	6	8	7	8	9	10	12	14	13	14	16	17	17
6	4	6	7	9	8	9	10	12	13	16	14	16	17	0	0
7	6	7	8	10	9	10	12	13	14	17	16	17	0	0	0
8	7	8	9	12	10	12	13	14	16	0	17	0	0	0	0
4	3	4	6	8	7	8	9	10	12	14	13	14	16	17	17
6	4	6	7	9	8	9	10	12	13	16	14	16	17	0	0
7	6	7	8	10	9	10	12	13	14	17	16	17	0	0	0
8	7	8	9	12	10	12	13	14	16	0	17	0	0	0	0
9	8	9	10	13	12	13	14	16	17	0	0	0	0	0	0
10	9	10	12	14	13	14	16	17	0	0	0	0	0	0	0
12	10	12	13	16	14	16	17	0	0	0	0	0	0	0	0
13	12	13	14	17	16	17	0	0	0	0	0	0	0	0	0
4	3	4	6	8	7	8	9	10	12	14	13	14	16	17	17
6	4	6	7	9	8	9	10	12	13	16	14	16	17	0	0
7	6	7	8	10	9	10	12	13	14	17	16	17	0	0	0
8	7	8	9	12	10	12	13	14	16	0	17	0	0	0	0
9	8	9	10	13	12	13	14	16	17	0	0	0	0	0	0
10	9	10	12	14	13	14	16	17	0	0	0	0	0	0	0
12	10	12	13	16	14	16	17	0	0	0	0	0	0	0	0
13	12	13	14	17	16	17	0	0	0	0	0	0	0	0	0
14	13	14	16	0	17	0	0	0	0	0	0	0	0	0	0
16	14	16	17	0	0	0	0	0	0	0	0	0	0	0	0
17	16	17	0	0	0	0	0	0	0	0	0	0	0	0	0
0	17	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Figure 2

Positions of the first characters of the phrases of the greedy parsing are 1, 4, 6, 7, 8, 9, 10, 12, 13, 14, 16.

Greedy Versus Optimal Parsing for Text Compression

Idea: Textual substitution algorithms typically employ *greedy parsing*, that is, at each stage, the longest possible string is taken as the next match. However, *optimal parsing* strategies may sometimes take a shorter match so that bigger savings can be achieved later.

Example:

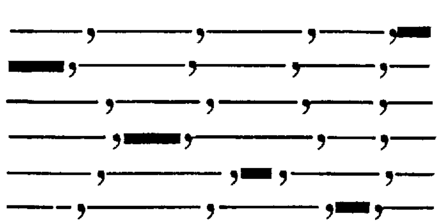
input string: $baba^k$

dictionary: $a, b, \{ba^i : 1 \leq i \leq k\}, bab$

optimal parsing: ba, ba^k

greedy parsing: $bab, a, a, \ldots a$

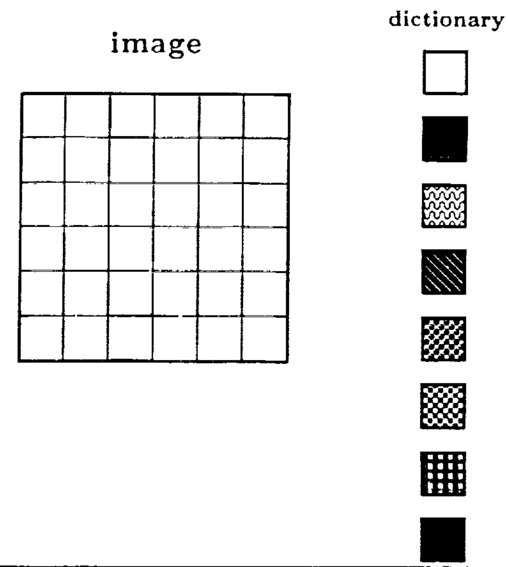
Theorem: The intersection of the i^{th} phrases of all optimal parsings is non empty, and hence there is a set of *canonical substrings* that identify the positions of optimal phrases



*** This is why optimal parsing can be done on-line and why parallel algorithms for greedy parsing can be generalized to optimal parsing.

Image Compression with Vector Quantization

IDEA: Map sub-arrays of pixels ("vectors") to the "closest" vector in a dictionary of vectors.



Key Results for Fixed-Sized (Trained) VQ

A complexity analysis of codebook design and search shows that optimal VQ as well as optimal k-means design is NP-hard.

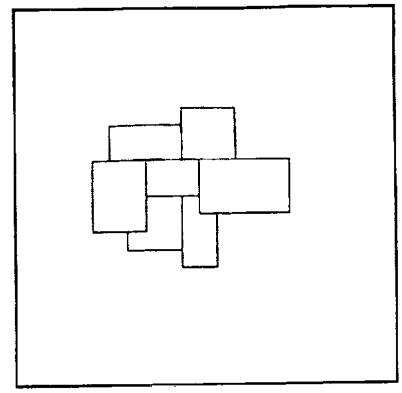
A complexity analysis of tree-structured VQ shows that finding optimal pruned trees subject to the leaf entropy or expected depth is NP-hard, whereas a polynomial-time algorithm is presented when the cost function is the number of leaves or maximum depth.

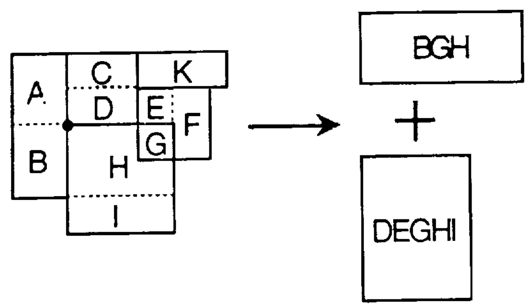
A new tree growing algorithm constructs trees that are well balanced among different costs; experimental results show it to achieve the compression performance of optimal pruning without the computational overhead.

Worst-case bounds on the performance of tree search versus optimal full search have been derived; efficient heuristics have been developed to significantly improve tree search performance.

Massively parallel algorithms have been developed to implement existing VQ algorithms as well as novel ones for design and codebook search.

On-Line Adaptive VQ





Experiments

Test images:

- **BrainMR:** 256x256, 8 bits/pixel brain image (mr)
- **BrainCAT:** 512x512, 8 bits/pixel brain image (cat-scan)
- **DonaldsonVilleLA4:** 512x512 NASA Satelite Image, band 4
- **DonaldsonVilleLA6:** 512x512 NASA Satelite Image, band 6
- **WomanHat:** 512x512, 8 bits/pixel grayscale photo
- **LivingRoom:** 512x512, 8 bits/pixel grayscale photo
- **FingerPrint:** 768x768, 8 bits/pixel FBI fingerprint image

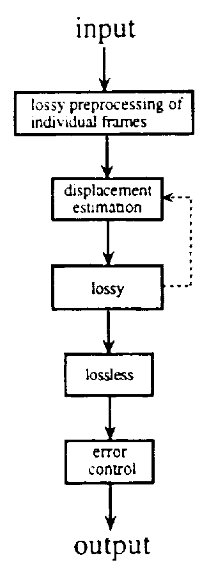
Visualization Tools:

- original image
- compressed image
- movie of adaptive growing process
- "checkerboard" display that maps the rectangles used
- median intensity display
- dictionary display
- error image

Results: Signal to noise ratios for a given compression ratio typically equal or better traditional fixed-size trained VQ. As can be seen from the slides of the decompressed data, resolution of edges is especially accurate (good for scientific and medical data).

NOTE: This is a huge success! The same adaptive method, with no prior knowledge of the data, can be used for diverse data sets while achieving the performance of trained methods.

Simplified Video Compression System



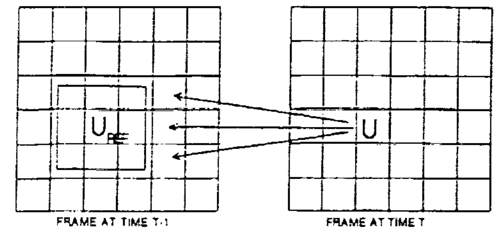
Displacement Estimation

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 .

Fixed Block Method (Jain and Jain [81])



$U = M \times N, \quad U_{REF} = (M+2p) \times (N+2p)$

Mean distortion function between U and U_{REF}:
$D(i,j) = \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N} g(u(m,n)-u_{ref}(m+i,n+j)), \quad -p \le i,j \le p$

DMD: direction of minimum distortion, for each block

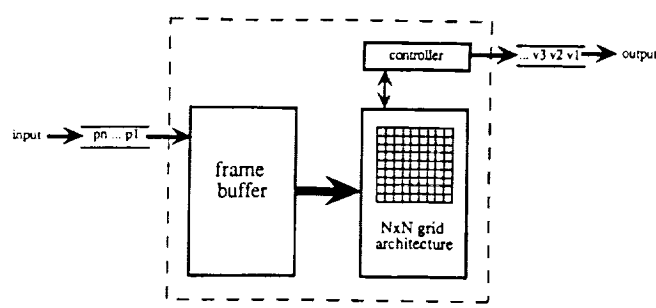
Finding DMD for a block = evaluation of D(i,j) in (2p+1) x (2p+1) positions

Hypothesis: data is such that D(i,j) monotonically increases as we move away from the DMD

Limitations of Traditional Displacement Estimation

- Monotonicity assumption.
- Fixed size blocks.
- Serial model of computation.

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

Encoder's Algorithm at time t

Phase 1 - all processors in parallel:

construct search area
compute DMD

Phase 2 - controller :

get DMD from each processor
compute superblock splits
output a list of all the splits
output DMD of each superblock

Phase 3 - controller :

construct new superblocks

Encoding Splits and Displacements

Splits : Send ID of the superblocks that need to be disoved.
Send list-of-splits.

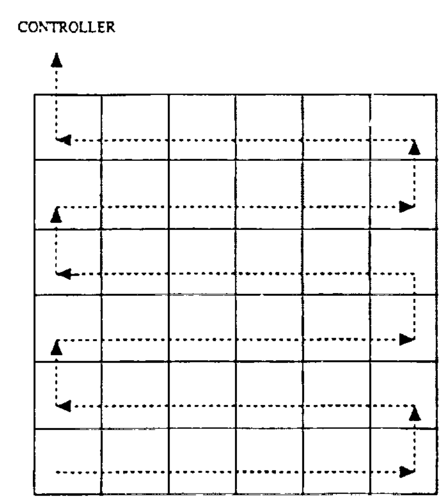
Displacements : Send one DMD for each superblock that has not been disolved and the DMD of the blocks for the superblocks that have been disolved.

Threshold Condition :

If *size_of_data* > *threshold*, controller disolves all the superblocks.

Note: Due to the threshold condition, never worse in terms of data sent than the fixed block approach.

Communication Between Controller and Processors



The communication between processors and controller is pipelined

Analysis of the Algorithm

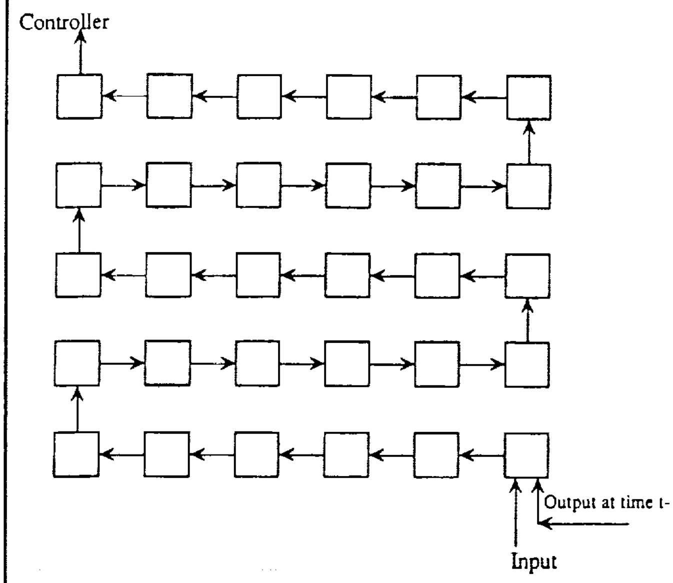
Time : O(n) (on-line algorithm).

Space : O(1) for processor
O(NxN) for controller.

Fidelity: no worse than fixed block method
(possibly better: no monotonicity assumption).

Amount of data sent:
no more than the fixed block algorithm (has the potential to represent "easy parts of a frame, e.g. background areas with single superblocks).

Implementing the Algorithm on a Pipe



Current Research

Lossless Compression:

- Polynomial-Time Algorithms for dynamic off-Line encoding, that are provably better than greedy, and are on-line decodable in sub-linear time.
- Parallel approximation algorithms for dynamic dictionary
- Sub-linear implementations on existing massively parallel machines (CM5, MASSPAR, etc.)
- A 2.5 billion bits per second systolic SONET board.

Adaptive Image Compression:

- Provable performance bounds (for both lossless and lossy).
- A better understanding of how quality is learned.
- New learning strategies.
- Implementations on existing massively parallel machines.
- Practical algorithms for hardware implementations.
- Visualization tools - ones that go beyond the tools we have developed thus far and and allow us to better understand "what is happening" to the data when it is compressed and to gain insights on improved growing / update strategies as well as distortion measures.

Video Compression:

- Poly-Log displacement estimation.
- Provable performance bounds.
- Further study of the complexity of various optimization problems arising in displacement estimation.
- ':nplementations on existing massively parallel machines.
- Practical algorithms for hardware implementations.
- Visualization tools.

PERFORMANCE AND SCALABILITY OF CLIENT-SERVER DATA BASE ARCHITECTURES

Dr. Alex Delis
University of Maryland

August 11, 1992

Performance and Scalability of Client–Server Database Architectures

Nick Roussopoulos
Alex Delis

Department of Computer Science
University of Maryland
College Park, MD 20742

Applied Information Systems Research Program (AISRP) Workshop II

August 11–13 1992
Boulder, Colorado

Today's Reality

- Autonomous database (lots of them)
- Overlapping and Inconsistent.
- Accessible through networks but not interoperable.
- Multiple sites, many server machines, thousands of workstations.
- Architectures: centralized or Client Server.
- Gateway access: one database at a time.
- Distributed heterogeneous DBMS?

Today's Needs

- Inter database querying.
- Download and Downsize.
- Inter database dependency tracking and change propagation.
- Version and Change Control.
- Architectures (distributed, pipelined, parallel).
- Interoperability of heterogeneous relational DBMSs.
- Multisite transaction management.

The Multidatabase Approach

- Databases are autonomous developed, managed and evolve independently.
- There is NO global schema.
- Each database has an *export schema* for allowing remote *access*.
- *Private schema* is for hiding portions of the database from remote users.
- No interference from remote transactions and no effect in the local applications.
- Inconsistency amongst multiple databases is an accepted fact of life.

Technology Trends

- Inexpensive and powerful (fast CPUs but relatively slow disks).
- Cumulative CPU reaches GIPS (mostly wasted).
- Current Software does not take advantage of workstation hardware.
- LANs are fast and they will soon be a lot faster.

Major Question:

How can we take advantage of these developments in order to develop DBMSs that offer fast response time and high throughput?

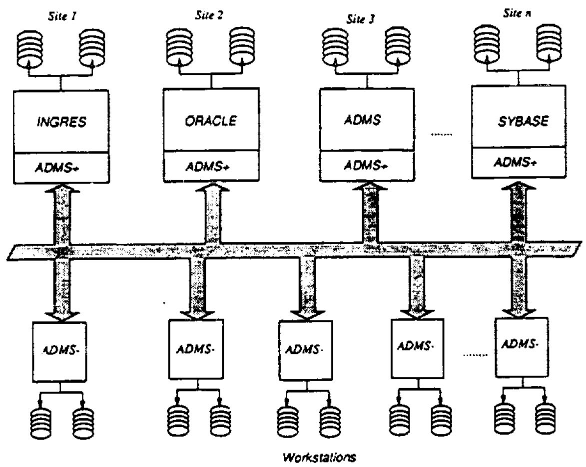
Motivation

- Execute Query Parsing and Optimization on the Workstations.
- Relieve the Server load by isolating intensive processing within the local workstation environment.
- Access the local disks in parallel.

Interoperability in DBMSs

- Gateway Interface (one at a time).
- Mutli window gateways (visual interface no exchange of messages).
- Interleaved environments (control and data exchange through messages.

ADMS±: Enhanced Client–Server Database Architecture with Incremental Gateways to Heterogeneous Relational DBMSs



- Every WS runs locally an ADMS- version of the ADMS.
- Every WS offers a serious disk capacity.
- The configuration follows the general **`Client-Server`** Model.

Main Features of the ADMS± System

EXPLOIT DATA CACHING:

- Every WS User may query server DBMS and choose to cache the data.
- There is gain when similar or related requests are being asked.

DIFFICULT POINT in CACHING:

- After a Server Update occurs, it may need to be propagated properly to all the pertinent clients.

How are Updates of Cached Data being carried out?

◇ **Update Logs [Differential Files]**

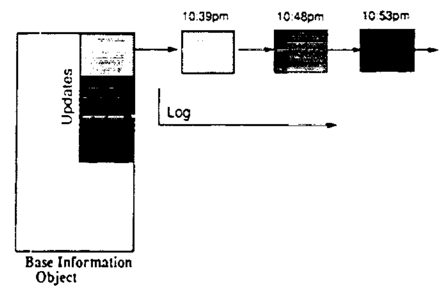


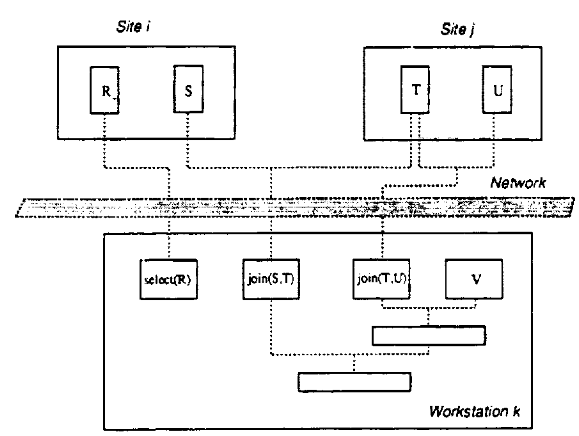
Figure 1: Log Overview

Main Features of the ADMS± System

- Every time a server relation is being queried then a bound between the server and the particular site is created (indicating parent child association).
- Updated portions of relations are Timestamped at the *Server* site.
- Every ADMS- keeps track of the last seen update timestamp of a server relation.
- From those relations having larger timestamps send to the appropriate ADMS- only the pertinent pieces of the Log ⟹ small increments are sent over the network.
- Queries and Updates are managed at the Server site utilizing a 2 ϕ Concurrency Control Protocol.

Global and Local Access Path Distribution

- Hybrid Access Paths: global and local.
- Global paths are serialized by the concurrency control module.
- Local paths run in parallel with other local paths.



ORIGINAL PAGE IS OF POOR QUALITY

Incremental Maintenance of Downloaded Data

- Efficient *differential* computation.
- Reduction in data transfer.

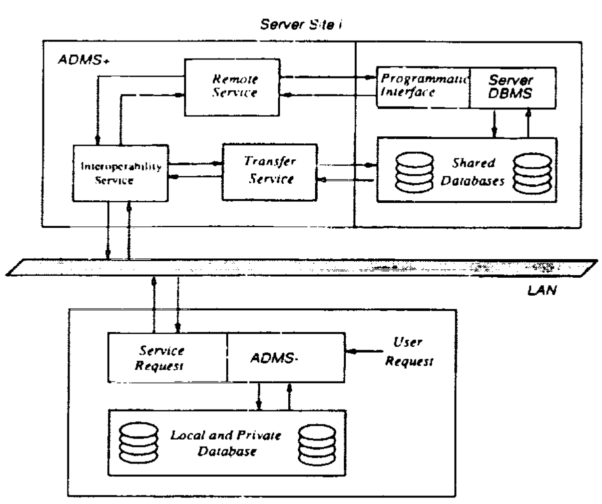
Deferred Update Strategies

- Lazy on demand update propagation.
- Periodic update propagation.

Advantages

- No broadcasting.
- Low communication overhead.
- Batched updates are optimized.

ADMS± Software Architecture



ADMS± Implementation

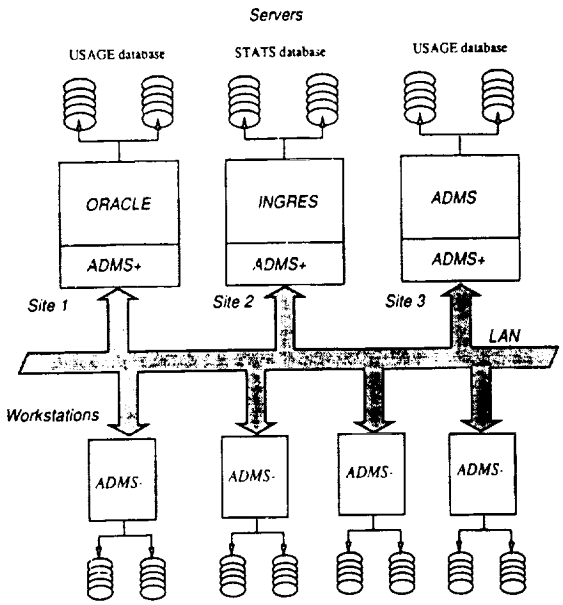
⊙ Prototype Platform: Unix on

- ⋆ Suns
- ⋆ DecStations
- ⋆ Vaxes

⊙ Gateways for:

- ⋆ Oracle
- ⋆ Sybase
- ⋆ Ingres

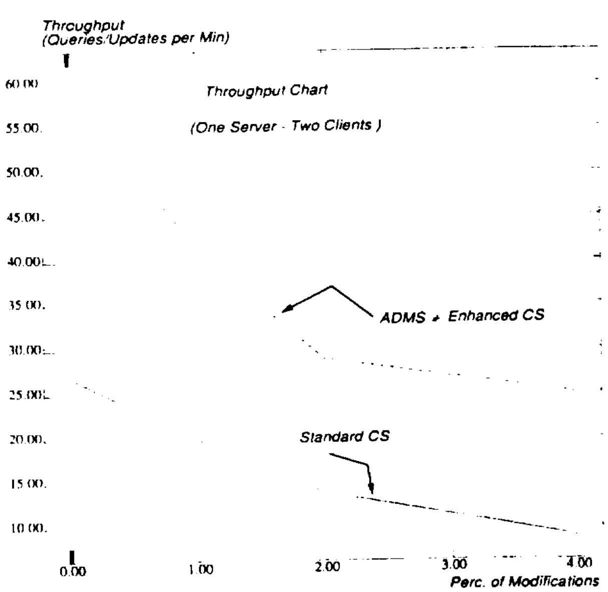
DEMO Configuration



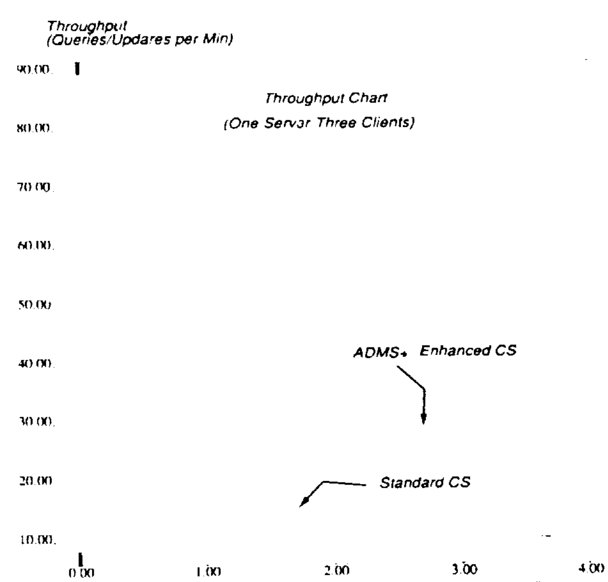
Prototype Behavior

- NFS and RPC behave well.
- Gateway Queries become a lot faster when most data is local
- Load on the servers dropped significantly.
- Parallel Access to the local data is the main contributor to the Query Throughput increase.

Combined Query Throughput



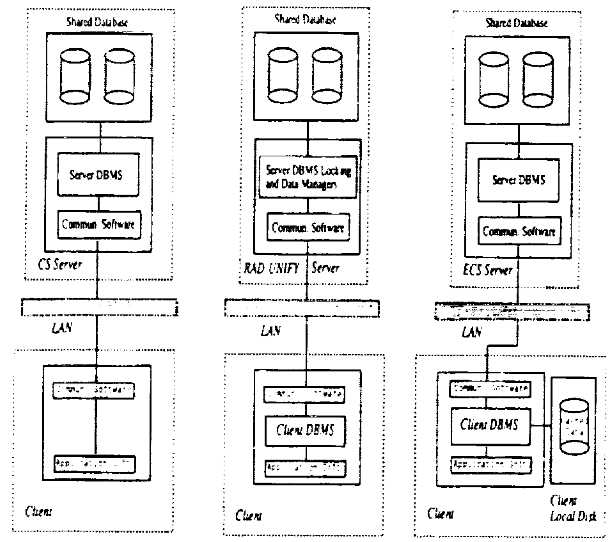
Combined Query Throughput



Capabilities

- Inter database queries and mixed breed views.
- Dynamic migration of data from the servers to the workstations.
- Incremental maintenance of downloaded data.
- Lazy no-broadcasting update strategies.
- Database servers are autonomous.
- Each workstation provides the *glue* for making the server database interoperable.
- Location transparency (but no database transparency).
- User perceives this as a really integrated distributed environment but does "feel" the distribution.

Comparison of DB Server Architectures



Standard Client-Server (CS) | Diskless Client-Server (RAD-UNIFY) | ADMS ± Enhanced Client-Server (ECS)

Characteristics

Standard Client-Server:

◇ **No distribution of data.**

- Overload
- Single Site

Diskless Client-Server:

[Rubenstein et al. SIGMOD 87]

◇ **Distribution of processing (better) but still**

- no site autonomy
- lots of net traffic
- concurrency overhead to all queries

ADMS± Enhanced Client-Server

[Roussopoulos et al. Computer 86]

◇ **Distribution of both processing and data**

- site autonomy (except for updates)
- minimal net traffic
- minimal overhead

Simulation Results

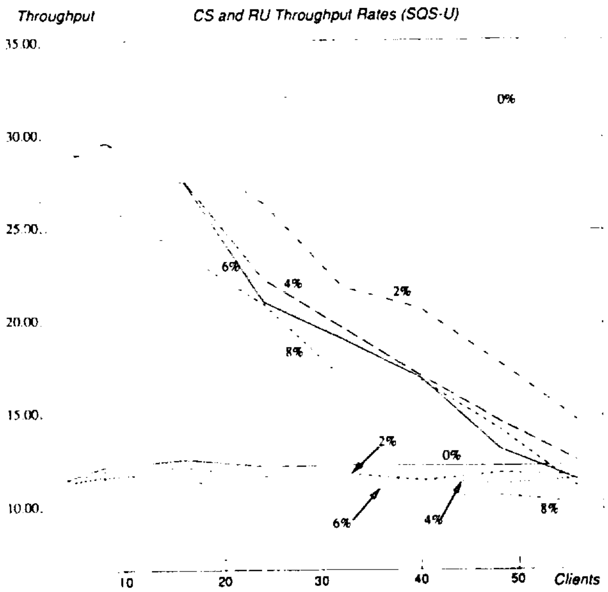
Parameters

- 1 Server n workstation clusters
- Query streams: three levels: light, medium, heavy I/O.
- CPU, I/O, Network Rate etc.

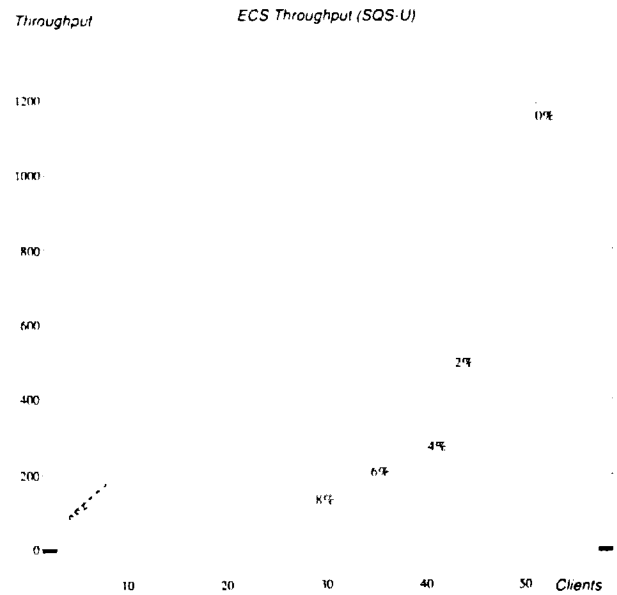
Metrics

- Query throughput and speedup
- Server I/O reduction
- Scale up

Simulation Results



Simulation Results



Summary

- Extensible, scalable and inexpensive architecture.
- Parallel access to multiple local copies.
- Incremental access methods for reducing data transfer and maintenance of downloaded data.
- Preserves database server autonomy.
- Glues multiple commercial DBMSs in a workstation environment.

Future Work on the ADMS± Architecture

- Gateway Query Optimization.
- Pipeline Algorithms for interdatabase queries (N way joins).
- Adaptive update propagation strategies.
- Multi site transaction management and recovery in autonomous databases.
- Experiment with increment updates of mirrored databases.
- Applicability of the same techniques in multi processor environment with or without shared memory.

A LAND-SURFACE TESTBED FOR THE EOS DATA INFORMATION SYSTEM (EOSDIS)

Dr. William Emery
Colorado Center for Astrodynamics Research

August 11, 1992

EOSDIS TESTBED SYSTEM

Provided by:
National Aeronautical and Space Administration (Nasa)
University of Colorado Colorado Center For Astrodynamic Research
University of California Santa Barbara
National Center for Atmospheric Research (Ncar)

Date:
August 10, 1992

Project Heads:
Bill Emery
Jeff Dozier
Paul Rotar

Project Design and Software Development Order

Devin Hooker
Tim Kelley

Project Design and Software Development Navorder

Tim Kelley

Navigation Design and Software Development

Dan Baldwin

Motifsho and Imagic Design and Software Development

Devin Hooker
Chuck Norris

This testbed data system provides advanced very high resolution radiometer (AVHRR) satellite images, image viewing, on line browse and manipulating software via the Internet. The software, Motifsho (for UNIX machines) and Imagic (for the Macintosh) were developed at the Colorado Center for Astrodynamic Research (CCAR). These display programs allow a user to look at an image, manipulate the histogram, crop, enhance via several modes, color, overlay, and animate the images. The major service of the testbed is to provide digital AVHRR data of the western U.S. to interested users connected to the Internet. The testbed system allows even the novice computer user to interface with the browse and ordering systems. The software is C and Motif X Windows programming, which allows users to easily interact with the testbed through the screen commands.

The raw satellite data and AVHRR images are stored on the National Center for Atmospheric Research (NCAR) Mass Storage Tape Drive System (MSS). All of the browse images are stored on the gateway workstation. Due to response time needs and processing constraints the testbed utilizes several programs to make the system successfully order and deliver the data. The data files that are stored on the MSS are accessed through the NCAR's Cray YMP super computer which retrieves the AVHRR images and sends the images to the gateway workstation.

The first step in the testbed system is for CCAR to track down link data from the NOAA weather satellites. The tracking system uses an ephemeris-driven orbital model to follow the satellite and start the telemetry ingest process. The telemetry stream transmitted from the satellite is received and recorded by the CCAR ground station. CCAR stores 3 to 4 satellite passes per day, seven days a week, each pass consists of 130 megabytes of data. The data is calibrated, channels separated, geo registered, navigated and formated. The second step of the testbed system is to send all of the raw and AVHRR images to the MSS. Once all of the data has been placed in the storage directories on the MSS, the gateway computer takes over.

The gateway computer polls the MSS everyday to update the catalog of all the AVHRR images and raw data. This catalog is then modified to interact with the other ordering and browse programs. Since the testbed programming is written is C and Motif X-Windows they are easily transportable to a wide array of computers on the Internet. Motif X-Windows provides the display capability and is a secure way to interact with all of the software and computer hardware, in that users are only given choices from a screen menu and will never get an entry prompt to break the secure environment of the window.

X Windows allows users to access Browse and Order, which are two separate programs that are accessed from the main programming on the gateway computer. Browse is a complex program that allows users to select an image from the current browse file list build the image, select the viewing size, map the image, overlay the map on the image, and send the image over the Internet from the host computer.

The browse image is a sample of the full image reduced by a factor of 4 in both the X and Y dimensions. The browse image is only useful in seeing the current weather over a specific region and area coverage from the satellite pass, while keeping the image small enough to be viewed over the Internet.

The order program works the same as Browse in that the user is given a window that interacts with a catalog file of images. The user has the option to order one or several images and have them returned to the ftp anonymous directory under their name. Order interacts with the MSS: it places the order, tracks the process number for the order, tracks the image order number, parse the mail from the MSS, and combines all of this to give the image its original name and order number. Order then ties the image to the user who ordered it, builds a file under the user's name in the ftp directory, and places the image in that directory. This process takes 5 minutes to complete, due mainly to the access time for the MSS. The current coverage area of a full AVHRR image consists of all of the west coast out to the grat plains using all five channels. The images are 2560(X) x 1540(Y) in pixel size.

Work is currently underway to upgrade the testbed interface to allow users the option of setting the parameters of the images they want. Users will pick the center latitude, longitude point, the type of projection (from a list of thirteen with the default being conic), the resolution, the viewing window size (512x512, 1024x1024 etc.), and the channels to be processed. Options also include the calculation of zenith angles, the processing of a high resolution overlay map, the development of a elevation mapped image, and the image bit size (8 bit or 2 byte). The new testbed system will work similar to the original except the new system will navigate and build the sub image requested by the user. Current tests show that this task can be accomplished within 5 minutes from the time the order is placed.

LIST OF PROMINENT USERS

UNITED STATES EDUCATIONAL FACILITIES

Purdue University
Northern Arizona University
University of New Mexico
Berkeley University
Tulane University
University of Utah
Penn State University
University of Arizona
University of Nebraska
University of Montana
Cornell University
Brigham Young University
Columbia
Colorado State University
University of Virginia
University of California, Santa Barbara
University of California, Davis
University of Texas
University of Oklahoma Storm Lab
Massachusetts Institute of Technology

UNITED STATES COMPANIES

Lockhead Corporation
McDonald Douglas corporation
Oxidental Corporation
MEC Analytical Systems
Lighting Sciences Inc
IBM
Hewlett Packard
Seamans

UNITED STATES GOVERNMENT AGENCIES

United States Navy
United States Air Force
NOAA
Pacific North Labs
U.S. Department of Agriculture
NASA
NASA JPL
U.S.G.S.
UCAR
SCRIPS
Woods Hole

FOREIGN EDUCATIONAL FACILITIES

University of Ottawa -- Canada
University of Toronto -- Canada
Osfold Regional College --Norway

FOREIGN GOVERNMENT AGENCIES

Department of Fisheries and Oceans -- Canada
United Kingdom Aero division

BREAK DOWN OF TEST BED GROWTH

Orders

Ordering has grown from zero order to as much as 400 orders per month. Currently, the test bed is going through transition to navorder. And the drop off can be traced to users waiting until navorder is on line and replaces order.

November 1991	1
December 1991	26
January 1992	90
February 1992	185
March 1992	420
April 1992	379
May 1992	400
June 1992	302
July 1992	279
Total Orders	**2,082**

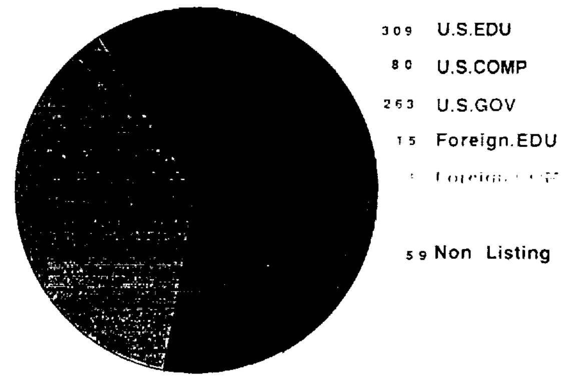
Users

Users are broken down into the following categories:
U.S. Education, U.S Companies, U.S. Government, Foreign Education, Foreign Companies, Foreign Government, and No Listing.

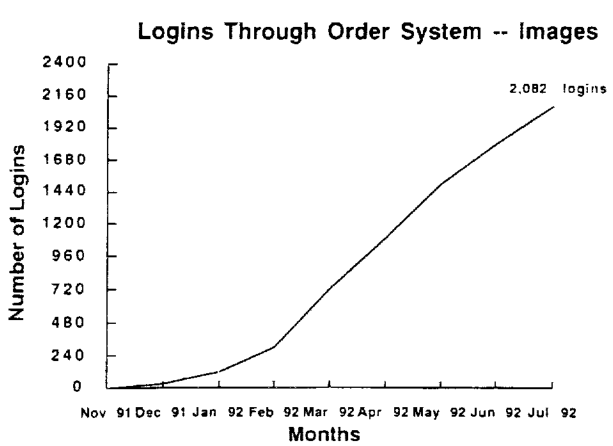
U.S. Education	309
U.S. Companies	80
U.S. Government	263
Foreign Education	15
Foreign Companies	4
Foreign Government	4
No Listing	59
Total Number of Logins	**734**

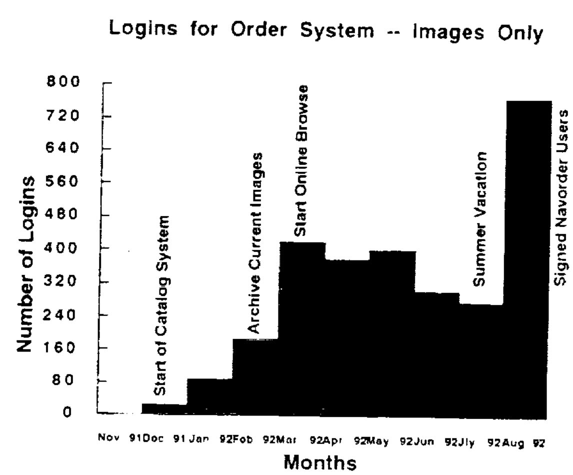
Graphs are provided for the data above.

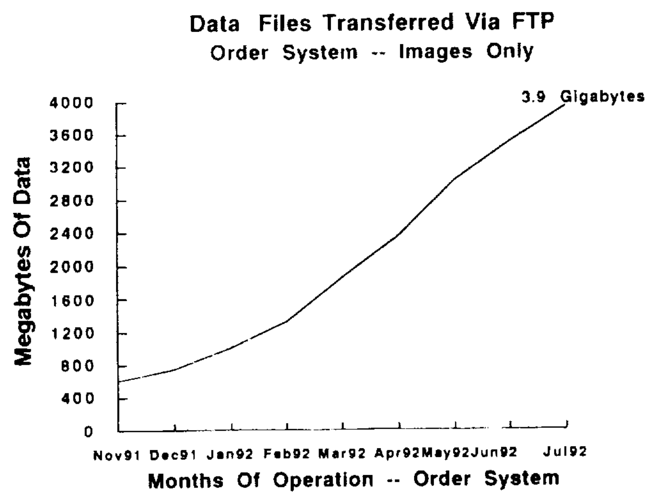
CHART OF ORDER PROGRAM USERS

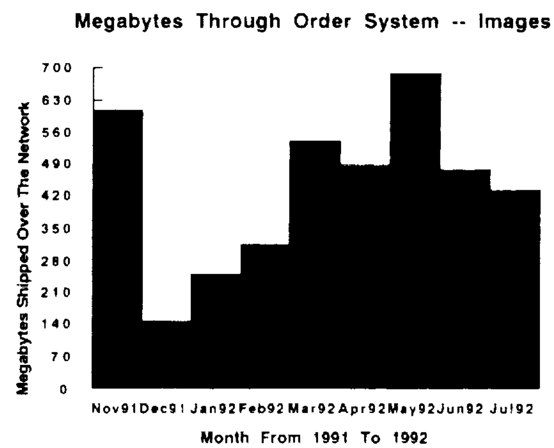


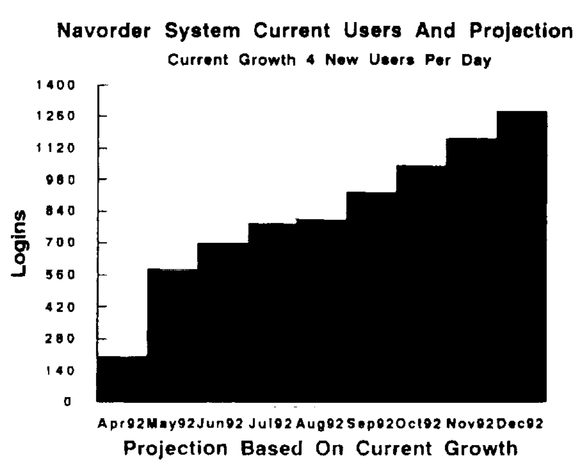
Total Agencies Using Order -- Images Only = 734

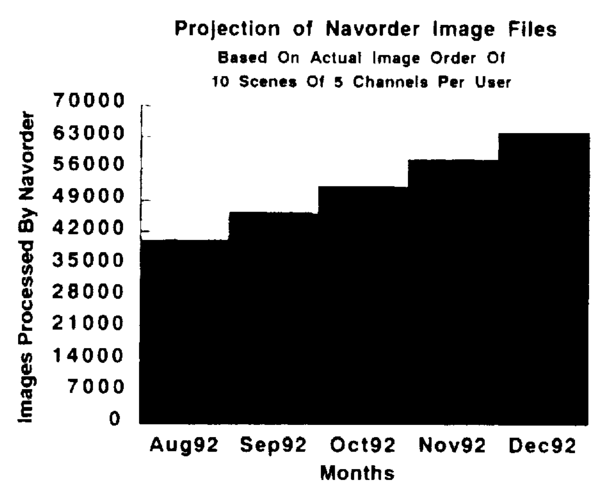





Data Files Transferred Via FTP
Order System -- Images Only
3.9 Gigabytes
Megabytes Of Data
4000
3600
3200
2800
2400
2000
1600
1200
800
400
0
Nov91 Dec91 Jan92 Feb92 Mar92 Apr92 May92 Jun92 Jul92
Months Of Operation -- Order System


Megabytes Through Order System -- Images
Megabytes Shipped Over The Network
700
630
560
490
420
350
280
210
140
70
0
Nov91 Dec91 Jan92 Feb92 Mar92 Apr92 May92 Jun92 Jul92
Month From 1991 To 1992


Navorder System Current Users And Projection
Current Growth 4 New Users Per Day
Logins
1400
1260
1120
980
840
700
560
420
280
140
0
Apr92 May92 Jun92 Jul92 Aug92 Sep92 Oct92 Nov92 Dec92
Projection Based On Current Growth


Projection of Navorder Image Files
Based On Actual Image Order Of
10 Scenes Of 5 Channels Per User
Images Processed By Navorder
70000
63000
56000
49000
42000
35000
28000
21000
14000
7000
0
Aug92 Sep92 Oct92 Nov92 Dec92
Months

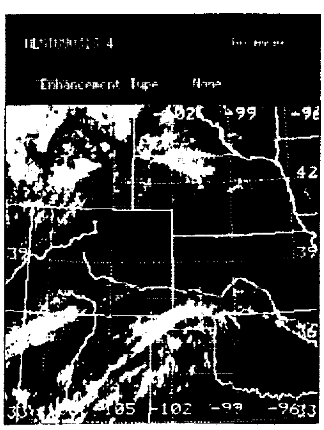
ORDER SYSTEM IMAGES ONLY

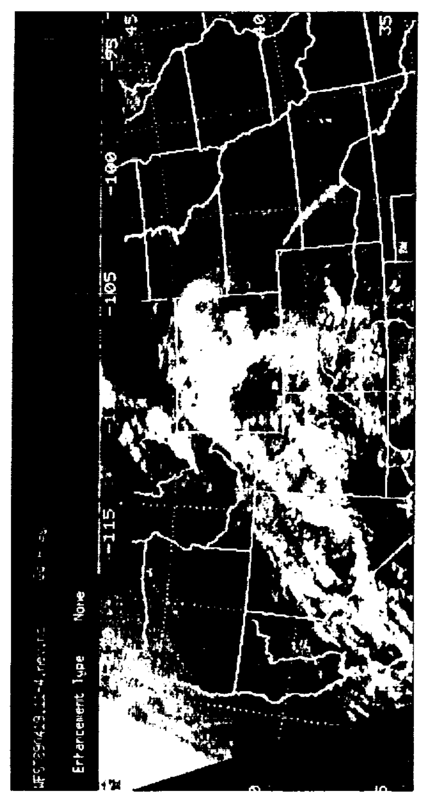
Examples Of Area Coverage And Ordering Window

Example One:
Image Catalog Window

Example Two:
First Area Coverage Browse

Example Three:
Latest Area Coverage Browse



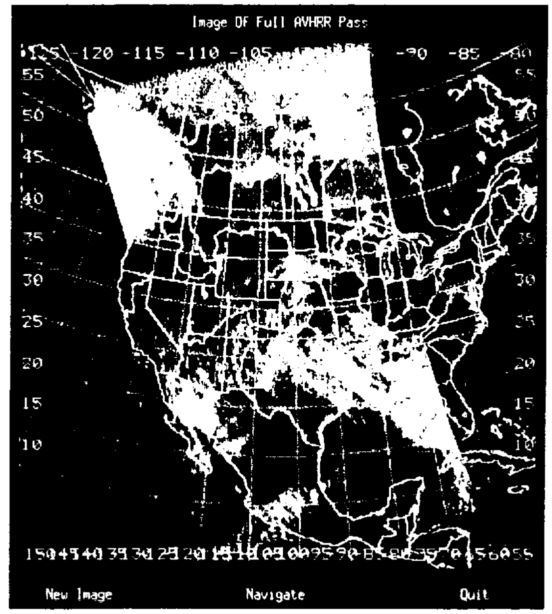


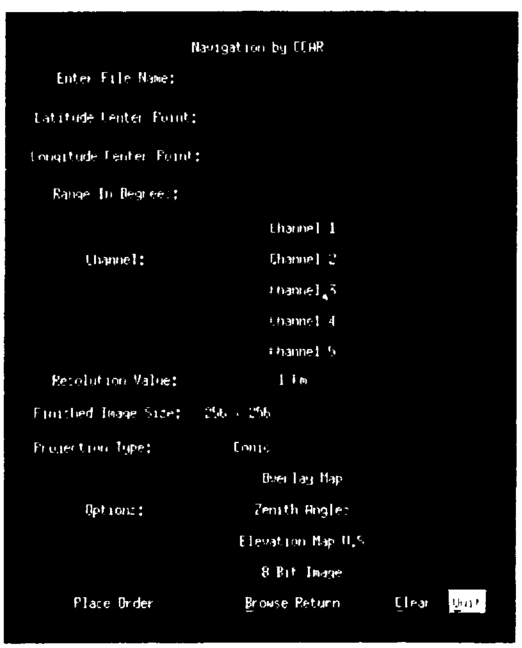
NAVORDER SYSTEM

Examples Of Area Coverage And Navigation Window

Example One:
Satellite Area Coverage Browse

Example Two:
Navigation Ordering Window





GEOGRAPHIC INFORMATION SYSTEM (GIS) FOR FUSION AND ANALYSIS OF HIGH-RESOLUTION REMOTE SENSING AND GROUND TRUTH DATA

Dr. Anthony Freeman
Jet Propulsion Laboratory

August 11, 1992

JPL

GEOGRAPHIC INFORMATION SYSTEM FOR FUSION AND ANALYSIS OF HIGH-RESOLUTION REMOTE SENSING AND GROUND TRUTH DATA

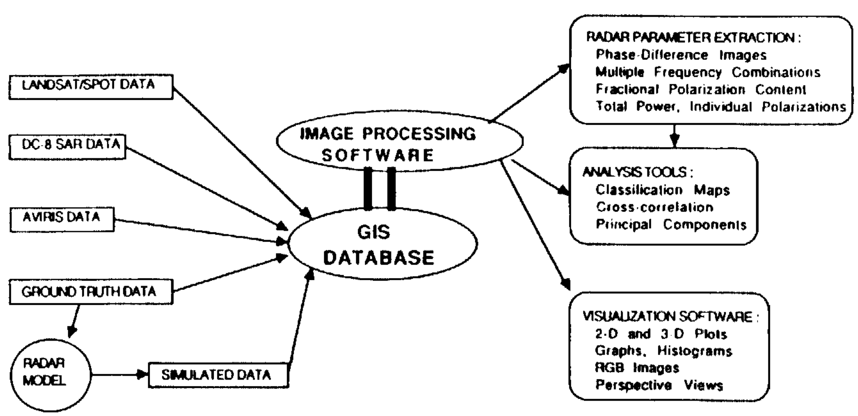
A. Freeman, J. Way and L. Norikane (JPL)

F. Leberl (Vexcel Corp.)

F. Davis and Y. Wang (UCSB)

JPL

Information Flow Diagram



JPL

Overview

- Three overflights of the Flevoland calibration/agricultural site were made by the JPL AIRSAR on July 3, July 12, and July 28, 1991.

- The modified VICAR/IBIS GIS was used to analyze these data. The following steps were taken:

[1] Generation of georeference image
[2] Coregistration of images to georeference image.
[3] Integration of data into georeference info file.
[4] Classification of data using selected features.
[5] Graphical analysis of data.

JPL

Modified VICAR/IBIS GIS

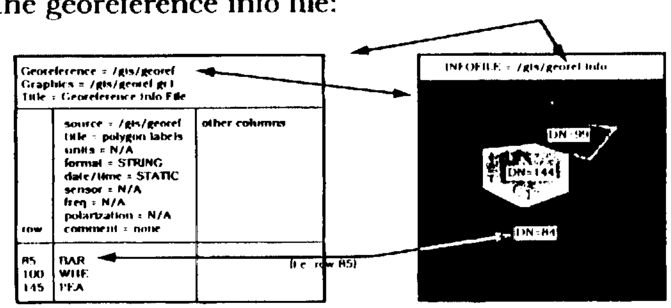
- Objectives
  - Be able to handle data in many different formats (vector, raster, tabular) and many different sources (models, radar images, ground truth surveys, optical images)
  - Link all data together through a georeference image
  - Track data in time, convert pixel values to "actual" values, plot graphs, generate training vectors for classification algorithms, compare actual and measured parameters
- Method
  - By using a new "info" file format, we can link many types of data to a georeference image. The "info" file allows tracking of data in time, conversions from pixel values to "actual" values, plotting, generation of training vectors for classification algorithms, and comparisons between actual and measured parameters.

JPL

- What is VICAR?
  - VICAR (Video Image Communication and Retrieval) is a set of programs and procedures designed to facilitate the acquisition, processing, and handling of digital image data. Its development began in 1966 at the Jet Propulsion Laboratory and is ongoing.
- What is VICAR/IBIS GIS?
  - VICAR/IBIS is a VICAR-based Geographical Information System (GIS). Its development began by Bryant at the Jet Propulsion Laboratory in 1970's. The VICAR/IBIS GIS requires that all image data be coregistered to a georeference image. Other types of non-image data are stored in columns of a "tabular" file which do not inherently contain any information about the data stored in each column.
- What is Modified VICAR/IBIS GIS?
  - The modified VICAR/IBIS GIS is an extension to the VICAR/IBIS GIS described above. By replacing the "tabular" file format with an "info" file format, we are able to satisfy the objectives of the GIS.

Generation of Georeference Image

- A polygon map was generated at TNO-FEL which overlayed the slant range projected July 3 data set.
- This map was used as the georeference image.
- Each polygon in the map had a unique ID number and a crop type. This information was stored in the georeference info file:



georeference info file
(e.g. "/gis/georef.info")

georeference image
(e.g. "/gis/georef")

JPL

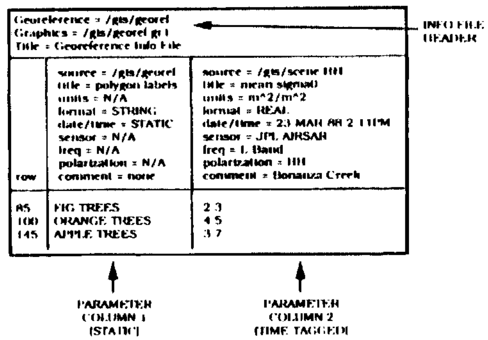
Image Registration

All images in the database must be registered to the georeference image. The registration process is composed of two steps: image rotation/scaling and image warping. The first step requires the selection of two tiepoints by hand. The second step requires the selection of multiple tiepoints selected either by hand or by using an automatic tiepoint generation program (L. Norikane).

An example of the results of using the automatic tiepoint generation program is shown on the next page.

JPL

- What is an "info" file?
  - The info file format is similar to the tabular file format in that data is stored in columns. Each column also contains a time tag, a link to its source file or source image, column format (real, integer, or string), and other information such as sensor, frequency, polarization, units, a title, and a comment. The file also contains a link to its associated georeference image.



Data Integration

- For each polygon in the georeference image, the following statistics can be calculated for each polygon in each data set at each frequency:

mean & std dev HH, HV, VV
mean HHVV* phase, HHVV* amplitude
mean correlation coefficient
mean odd bounce, double bounce, and volume scattering contributions (A. Freeman & S. Durden)

- Mean incidence angles are also calculated for each polygon in each data set.

- These statistics are saved in columns of the georeference info file.

JPL

The Info File and Data Integration

Different types of data can be integrated into the GIS database via the info file.

For remotely sensed data such as polarimetric SAR data, different images can be synthesized. The images created are byte images, but each pixel value represents a particular radar backscatter measurement. To retain this information, a DN CONVERSION column is created in each image's associated info file.

For each polygon in the georeference image, the mean and standard deviation of the "actual" values of the pixels which fall within each polygon can be stored in the georeference info file.

Spreadsheet data can be integrated into the GIS via the georeference info file. The spreadsheet must have a column containing a link to the georeference image, usually via a polygon label.

Other data can be entered manually into the georeference info file.

JPL

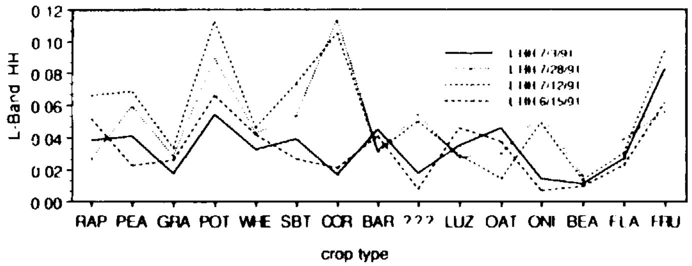
The Info File and Analysis

Info files are essential to the analysis process. They are required for making graphical plots to visualize temporal changes or possible correlations between different parameters. Selected rows and columns of the info file can also be used as training vectors for supervised classification algorithms or as inputs to modeling algorithms.

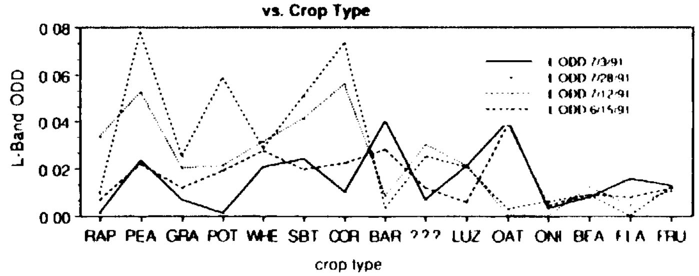
The next page contains an example of a graphical plot generated from selected rows and columns of a geoereference info file.

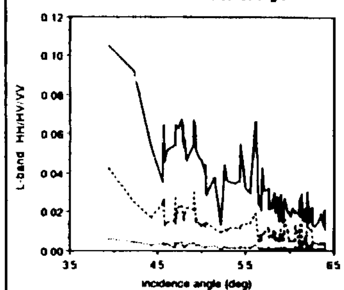
The last page contains an example of an actual crop map and a crop map generated by a supervised Bayes classifier using training vectors from the georeference info file (van den Brock).

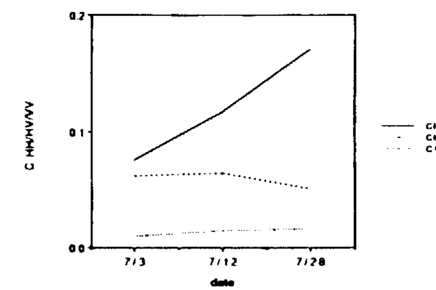
Temporal Variation of L-Band HH vs. Crop Type

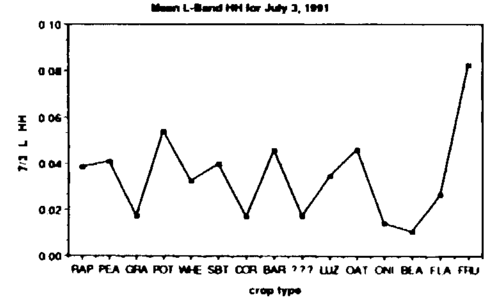


Temporal Variation of L-Band Odd Bounce Scattering Component vs. Crop Type









Confusion matrices

% classified as...

actual crop type

	RAP	PEA	GRA	POT	WHE	SBT	COR	BAR	???	LUZ	OAT	ONI	BEA	FLA	FRU
RAP	97.9	0.0	0.0	1.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2
PEA	0.2	96.2	0.1	1.8	0.0	1.3	0.0	0.2	0.0	0.1	0.1	0.0	0.0	0.0	0.1
GRA	0.1	0.0	81.0	3.0	1.3	3.9	1.0	0.0	0.3	5.8	0.3	2.7	0.3	0.0	0.2
POT	5.1	0.1	1.0	59.0	3.2	0.3	0.0	1.4	0.0	0.0	28.9	0.0	0.1	0.0	0.8
WHE	0.1	0.0	0.0	0.3	96.7	1.3	0.0	1.3	0.0	0.1	0.0	0.0	0.0	0.0	0.0
SBT	1.9	0.0	0.4	0.6	0.1	96.3	0.0	0.0	0.1	0.0	0.0	0.5	0.0	0.0	0.1
COR	0.0	0.0	4.2	0.0	3.4	1.4	79.1	0.0	0.0	11.3	0.0	0.3	0.3	0.0	0.0
BAR	0.2	0.1	0.0	0.2	11.7	0.2	0.0	47.6	0.4	0.1	0.4	0.0	3.9	0.4	0.7
???	5.8	0.0	0.5	0.0	0.0	5.8	0.0	0.0	86.8	0.5	0.0	0.5	0.0	0.0	0.0
LUZ	0.0	0.4	1.3	0.1	2.7	2.0	6.2	0.0	0.0	86.5	0.0	0.5	0.3	0.0	0.0
OAT	0.0	0.8	0.4	8.8	13.1	2.4	0.0	4.8	0.0	0.1	68.2	0.0	0.4	0.0	0.0
ONI	1.1	0.0	9.3	0.0	0.0	5.5	1.9	0.0	0.7	3.1	0.0	78.0	0.0	0.0	0.4
BEA	0.0	0.0	0.5	0.1	4.8	1.0	0.1	2.1	0.1	1.6	1.2	0.0	88.1	0.6	0.0
FLA	0.2	0.0	1.4	0.9	0.5	1.4	0.0	0.2	0.2	0.1	0.0	0.0	0.4	94.8	0.0
FRU	0.0	0.0	0.0	0.0	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	99.7

(features used. July 3 C/L/P HH/HV/VV)

% classified as...

actual crop type

	RAP	PEA	GRA	POT	WHE	SBT	COR	BAR	???	LUZ	OAT	ONI	BEA	FLA	FRU
RAP	99.2	0.0	0.0	0.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.3
PEA	0.0	93.7	0.0	0.4	0.6	2.8	0.0	1.7	0.0	0.0	0.0	0.0	0.0	0.0	0.8
GRA	0.9	0.1	83.9	0.0	10.2	0.1	0.0	0.0	0.0	4.8	0.0	0.0	0.0	0.0	0.0
POT	0.0	0.0	0.0	84.9	0.0	0.3	0.0	1.8	0.0	0.0	12.9	0.0	0.0	0.1	0.1
WHE	0.0	0.1	0.2	0.0	97.9	0.6	0.0	0.9	0.0	0.0	0.0	0.0	0.0	0.0	0.2
SBT	0.1	0.1	0.1	0.0	0.4	98.6	0.0	0.4	0.0	0.0	0.0	0.0	0.0	0.0	0.4
COR	0.0	0.0	0.0	0.0	0.0	2.8	94.9	0.0	0.0	0.0	0.0	2.3	0.0	0.0	0.0
BAR	0.1	0.0	0.2	1.7	0.1	0.4	0.0	89.2	0.0	0.0	0.0	0.0	7.1	0.0	1.1
???	0.0	0.5	0.0	0.0	0.0	5.8	0.5	5.3	86.2	0.0	0.0	0.5	0.0	0.0	1.1
LUZ	1.3	0.0	0.2	0.2	0.0	0.8	0.0	0.1	0.0	97.4	0.0	0.0	0.1	0.0	0.0
OAT	0.0	0.0	0.1	20.6	0.3	1.1	0.0	5.8	0.6	0.0	71.3	0.0	0.1	0.0	0.0
ONI	0.0	0.0	0.0	0.0	0.0	3.5	6.0	0.0	0.0	0.0	0.0	89.5	0.0	0.0	1.1
BEA	0.1	0.1	0.6	1.4	6.3	0.1	0.0	11.1	0.0	0.1	0.0	0.0	79.4	0.6	0.0
FLA	0.0	0.0	0.0	3.0	0.0	0.7	0.0	2.4	0.1	0.0	0.1	0.1	1.4	92.3	0.1
FRU	0.0	0.0	0.0	0.0	0.0	0.8	0.0	0.3	0.0	0.0	0.0	0.0	0.0	0.0	99.0

(features used: July 3/12/28 L HH/HV/VV)

JPL

Data Reduction

The technique of principal components analysis has been explored as a method of data reduction for coregistered images. In deriving the principal components images from a set of coregistered images, the "actual" values are used rather than pixel values. The resulting principal components images, however, do not necessarily retain a meaningful conversion from pixel value back to some "actual" value. Thus these images do not have info files associated with them.

PC 1 PC 2 PC 3

Principal Components of SPOT Images of Bonanza Creek, Alaska

JPL

PROGRESS REPORT

- A working GIS/Image Processing System has now been integrated
- GIS has been exercised using multi-temporal data from the boreal forest test site and an agricultural site
- Classified rain forest site using radar data alone
- Integrated UCSB continuous and discrete canopy models and conducted a sensitivity analysis of the models (UCSB)
- Installed 'Light-table' - S/W tool for interacting with very large images (Vexcel)
- Developed a model which estimates scattering mechanisms from radar data
- Developed MACsigma0 S/W for release to AIRSAR data users

WHAT'S NEXT?

- Ground Truth data entry and correlation with image data
- Complete integration of UCSB model with GIS
- Analysis of boreal forest data
- New technique for feature selection/classification
- Rain forest site vegetation cover map
- Varying terrain height data registration
- SIR-C data analysis tools?

ENVISION: AN ANALYSIS AND DISPLAY SYSTEM FOR LARGE GEOPHYSICAL DATA SETS

Dr. Kenneth P. Bowman
Texas A & M University

August 11, 1992

ENVISION

An Analysis and Display System for Large Geophysical Data Sets

Principle Investigators

Kenneth P. Bowman
Climate Systems Research Program
Department of Meteorology
Texas A&M University

John E. Walsh
Department of Atmospheric Sciences
University of Illinois at Urbana–Champaign

Robert B. Wilhelmson
National Center for Supercomputing Applications
Department of Atmospheric Sciences
University of Illinois at Urbana–Champaign

Designers/Programmers

Keith R. Searight
David P. Wojtowicz
Department of Atmospheric Sciences
University of Illinois at Urbana–Champaign

Envision consists of:

- a metadata browser and editor
- a data management system
- a set of links to feed data to existing visualization tools
- a set of custom designed visualization, analysis and data manipulation tools

Envision is *Not*:

- a database system, relational or otherwise
- a new file format
- a toolkit or library
- a specific visualization tool

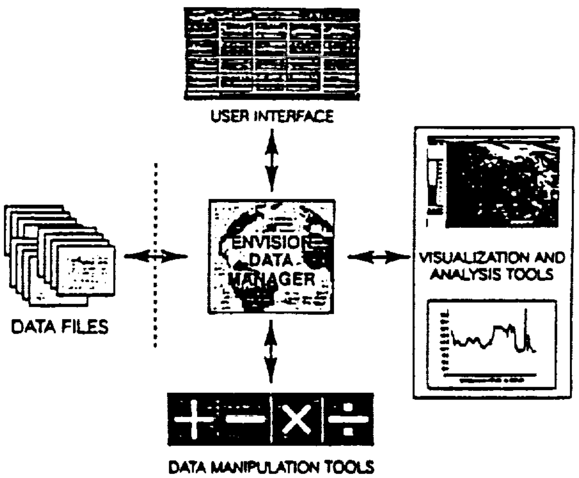
ENVISION

Data type requirements for Envision:

- regular nD grids
- grids may contain missing data or undefined regions
- these grids currently must be stored in netCDF files

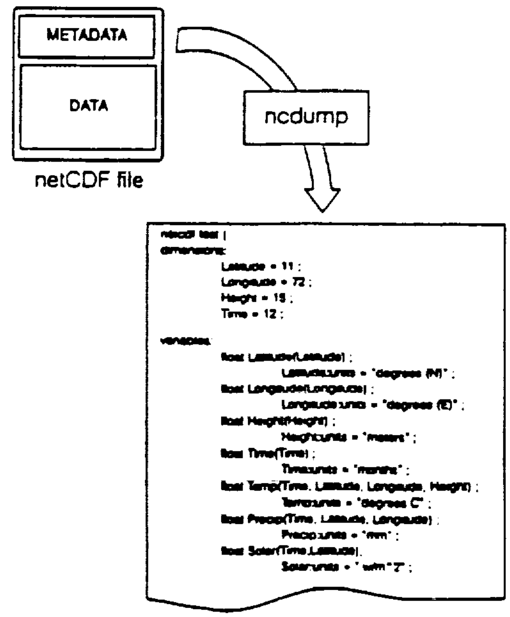
ENVISION

Envision system layout:



ENVISION

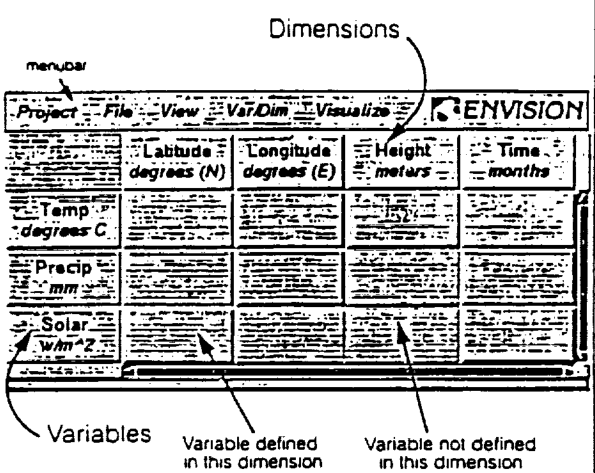
The usual way to see what's inside a netCDF file:



ENVISION

Envision as a Metadata Browser:

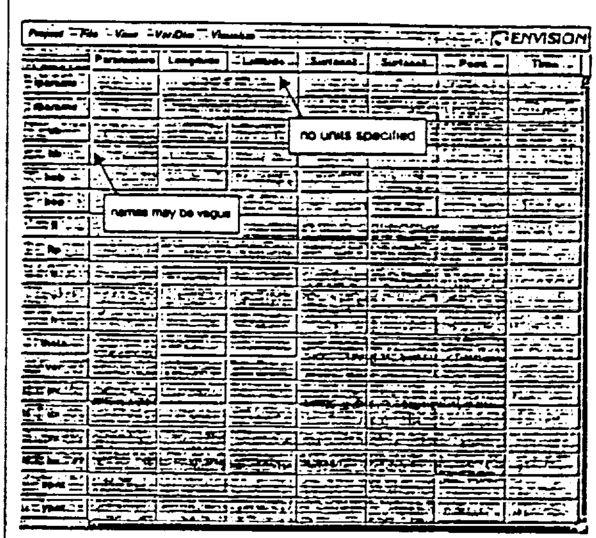
The Envision Table Display / User Interface



ENVISION

Envision as a Metadata Editor:

Typical file with only the minimum required metadata:



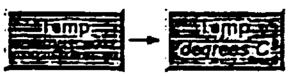
ENVISION

Envision as a Metadata Editor:

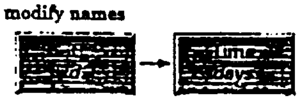
Allows the user to modify, augment or delete file metadata.

Examples:

- add units



- modify names



ENVISION

Envision as a Metadata Editor:

It is not necessary to write these changes to the file.

This is desirable because:

- the file may not be writeable (CDROM, user doesn't own file, etc.)
- netCDF library rewrites entire file if metadata is expanded
- user may not want changes to be permanent

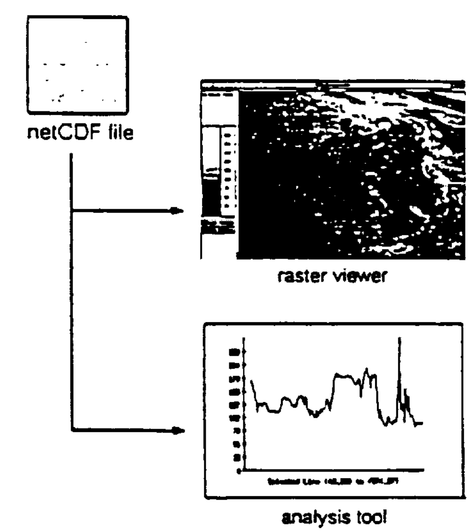
Data is stored externally by Envision.

However, changes to metadata may be incorporated into actual data files at any time user desires.

ENVISION

Envision as a Data Manager:

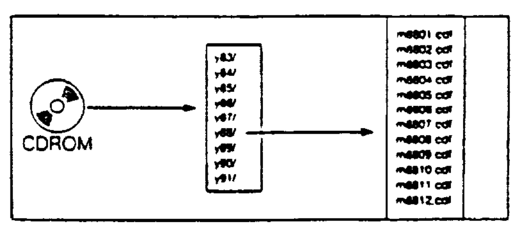
Life is simple if all your data is in one netCDF file.



ENVISION

Envision as a Data Manager:

Often however, a dataset consists of more than one file.

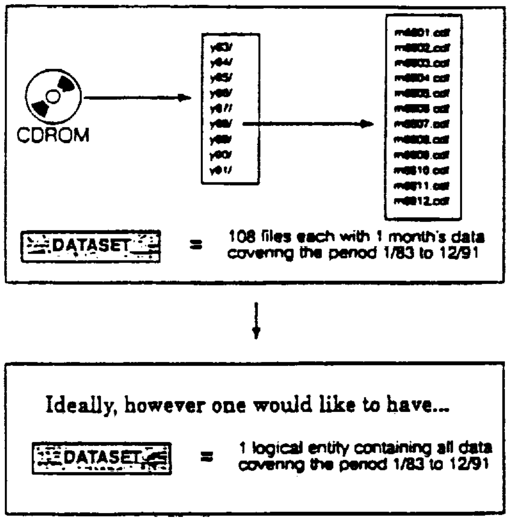


This makes it difficult to deal with the dataset as a whole

- files must be processed individually
- the user must keep track of which file the desired data is located in

ENVISION

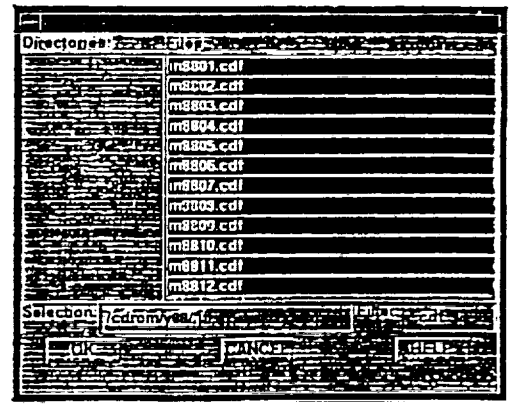
Envision as a Data Manager:



ENVISION

Envision as a Data Manager:

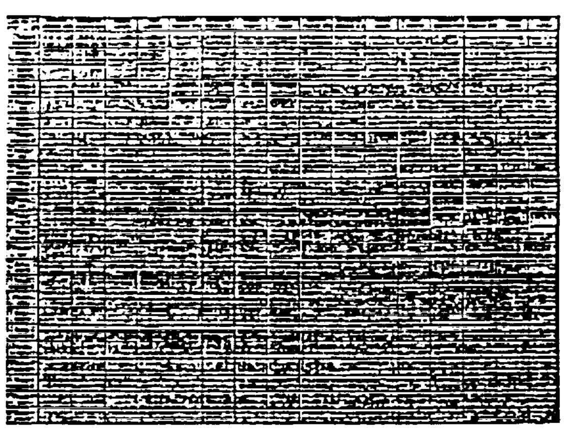
Envision can deal with data contained in multiple files.



Multiple files being selected to be added to the table display.

ENVISION

Envision as a Data Manager:



After loading multiple files, the variables and dimensions contained within these files are displayed together.

ENVISION

Envision as a Data Manager:

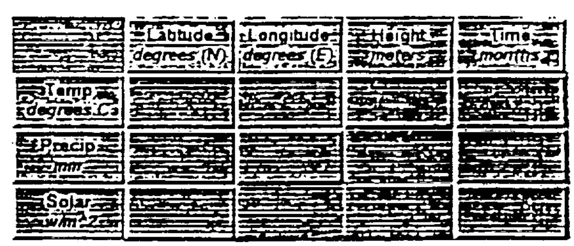


Table display representing the metadata within the files after merging like variables and dimensions.

ENVISION

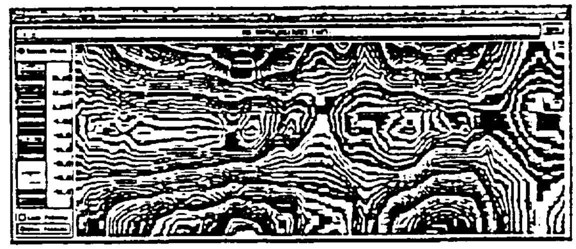
Envision as a Data Manager:

Summary of data management capabilities:

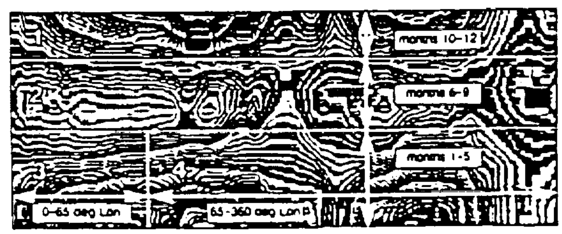
- manages relationships between files
- provides transparent access as a single entity to a dataset consisting of multiple files
- delivers arbitrary 1,2,3D "slabs" of data from a dataset to visualization, analysis and manipulation tools
- special emphasis on time problems, etc.

ENVISION

Envision as a Data Manager:

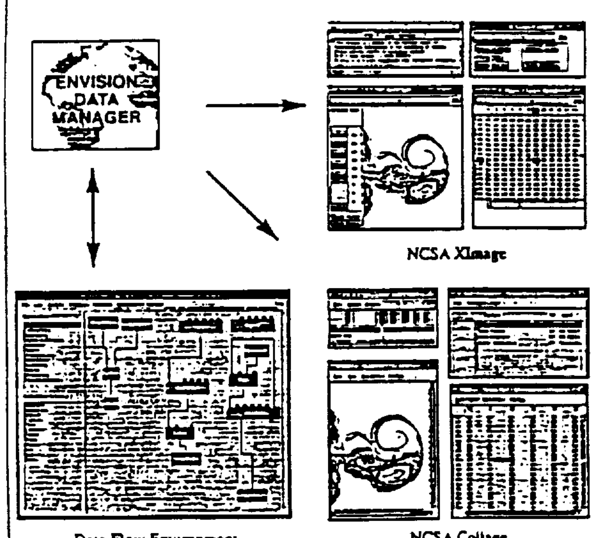


Example: visualization of a range of data that spans four files.



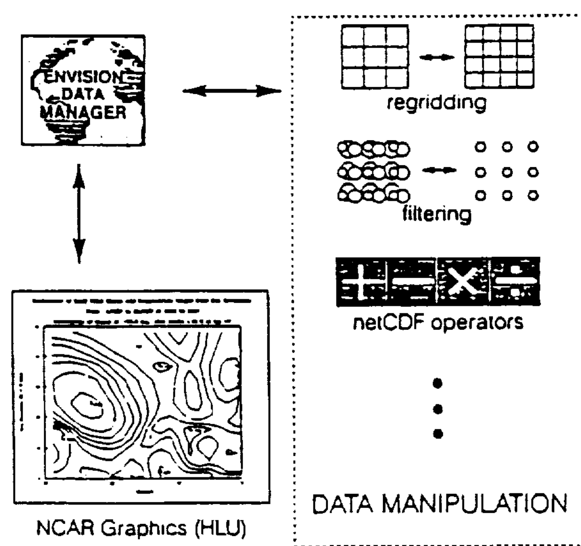
ENVISION

Envision as a data source for existing visualization tools:



ENVISION

Envision as a data source for custom visualization, analysis, and data manipulation tools:



ENVISION

Additional features of Envision:

- distributed processing
- context help
- multiple linked user interfaces
- demo capability

ENVISION

Current Status:

- Currently 11 months into project
- Data management code and network interface mostly complete
- User interface about 50% complete
- Links to *XImage* and *Collage* working

Goals:

Release in late 1992 of a system with:

- Envision data management facility
- Envision interface
- Customized connections to:
  - NCSA *XImage*
  - NCSA *Collage*
  - Some *NCAR Graphics* utilities (contours)

ENVISION

A GLOBAL SATELLITE DATA ACQUISITION AND ANALYSIS SYSTEM TO SUPPORT HYDROLOGICAL MODELING AND REGIONAL CLIMATIC CHANGE IMPACT STUDIES

Dr. Fran Stetina
NASA/Goddard Space Flight Center

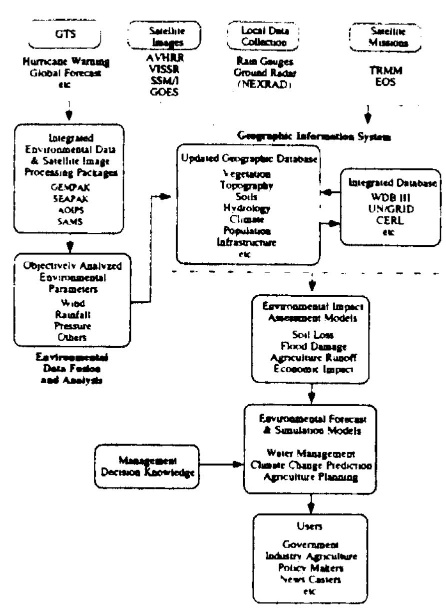
August 11, 1992

TITLE: "A Global Satellite Data Acquisition and Analysis System to Support Hydrological Modeling and Regional Climatic Change Impact Studies"

Abstract:

This is a proposal to develop an uniform global environmental data gathering and distribution system to support the calibration and validation of remotely sensed data. This system consists of state-of-the-art graphics and visualization techniques, simulations models, database management and expert systems for conducting hydrological and other global change 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 TRMM. This will be a joint NASA-U.S. Army Corp of Engineers-FEMA-U.S. Bureau of Reclamation-University-Industry project.

AUTHORS: Charles Vermillion, Fran Stetina
NASA/Goddard Space Flight Center
John M. Hill, Houston Advanced Research Center
Craig Fischenich, U.S. Army Corp of Engineers
Charles Jones, Texas Agricultural Experiment Station
Daniel Cotter, Federal Emergency Management Agency
Federal Insurance Agency
Bernie Silverman, U.S. Bureau of Reclamation,
Thailand Royal Rainmaking Research & Dev. Institute



Integrated System Design of Environmental Impact Assessment Modeling System.

Figure 2: Operation of the reformat compiler

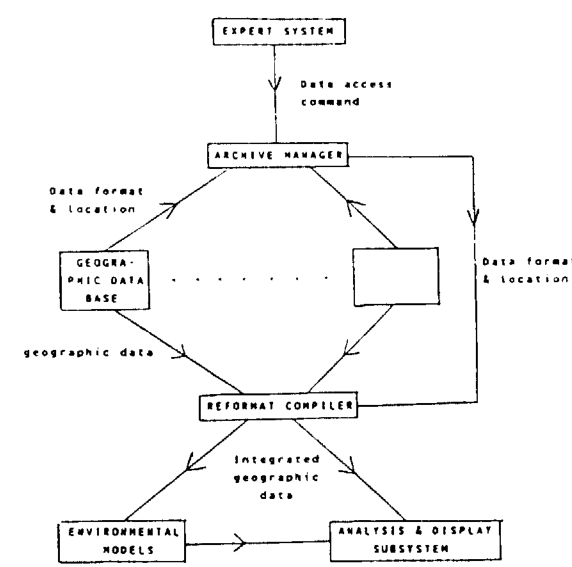
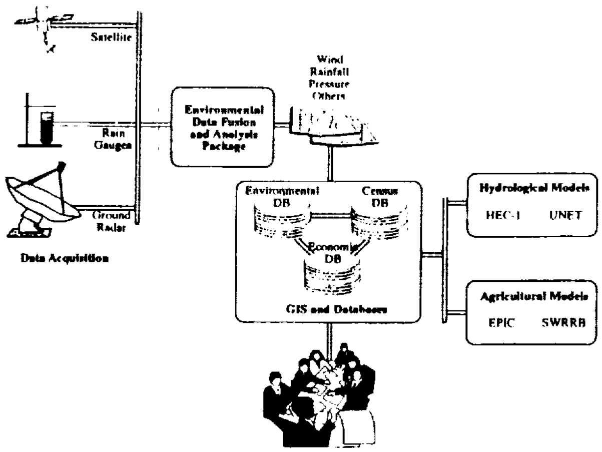
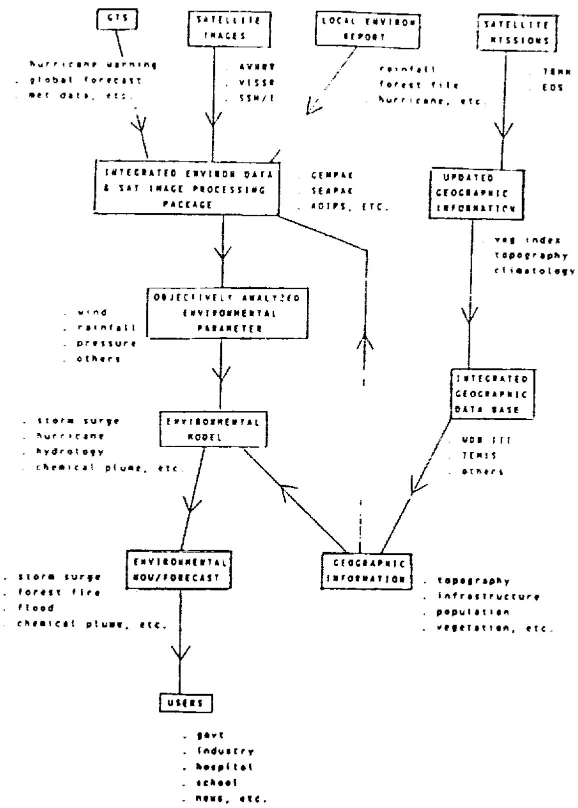


FIGURE 3



Environmental Impact Assessment Modeling System.

FIGURE 4- TOP LEVEL DESIGN OF EIAMS



1. INTRODUCTION

The objective of this proposal is to integrate a real time satellite acquisition and data analysis system with environmental algorithms and models to facilitate data acquisition for environmental monitoring and management. Both the satellite system and environmental models are in existence. We will develop an integrated computer environment for these packages, and make this integrated system transportable to most UNIX platforms; and make the system available to the science and application community to conduct global change impact studies.

Most software systems for analysis of data and images from various satellites exist independently of each other; and not much effort has been devoted to integrating them into uniform systems. Such lack of uniformity poses a barrier to users who often have to search for ad hoc hardware and software in analyzing environmental information from a multitude of satellites.

We will develop an Environmental Impact Assessment and Modeling System (EIAMS) as a framework for data gathering, analysis and distribution to facilitate real-time environmental monitoring and hydrological modeling. Such a system is required by earth resource scientists and managers to conduct regional climate analysis and impact analysis studies.

Currently, components of EIAMS exist. These components include a set of tools which can be used to analyze data and images from multiple satellites (or instruments) DMSP, TIROS-AVHRR, GOES, METEOSAT, GMS and INSAT (including SSM/I and TOVS). A package for analyzing hydro-meteorological data, GEMPAK, has also been integrated under this software system. Potentially this system can accommodate data from other instruments and satellites. Funded under this proposal, the next stage will be to integrate satellite rain-rate retrieval algorithms, hydrologic and agriculture models, and a geographic information system (GIS), all under one computing environment. The GIS and the hydrologic models are provided by the U.S. Army Corps of Engineers (COE), and the agricultural models by Houston Advanced Research Center (HARC).

The uniqueness of the proposed EIAMS is that it integrates several existing regional data bases, environmental monitoring systems and models into one computing environment. Numerous public domain software systems will be integrated and form the core of the system.

Large geographic data bases exist and will be extended by planned satellite monitoring missions such as Tropical Rainfall Measurement Mission (TRMM), and direct broadcast data from SEAWIFS and the Earth Observing System (EOS). EIAMS will provide a framework to integrate and facilitate utilization of the new data sets from these future NASA satellite missions.

The primary focus of this project would be to integrate the above set of analysis tools and models under one system, and to provide a graphical user interface. The system will be data-driven, instead of image-driven. Hence it will enable the scientists to perform scientific analysis as well as image processing.

To test this system, several closely related applications would be proposed. For example, one of them is the analysis of SSM/I and GMS data to derive a rain-rate diurnal cycle over the Western Pacific Ocean to support TOGA-COARE and over Thailand to support TRMM and Royal Rainmaking Research Institute. The Principal Investigators of this proposal have assisted the USAF in developing a real-time satellite data ingest and analysis system at the Joint Typhoon Warning Center in Guam. This satellite ground system can receive GMS, TIROS and DMSP. The proposed software will be tested with data from this satellite ground system.

Further system tests will be conducted to evaluate the system as regards meeting the users requirements of both applications and research scientists. These system evaluation tests will be conducted to support co-investigators from U.S. Army Corp of Engineers and Federal Emergency Management Agency, the Bureau of Reclamation, Royal Rainmaking Research and Development Institute of Thailand and the Texas Agricultural Experiment Station.

In summary, the critical development areas of this system are:

1) The system will have data input from a number of sources. Hence a data management software is needed for reformatting data inputs and outputs and data transactions between processes.

2) The system will have a set of basic analysis tools such as image analysis and statistical analysis. The system will also possess sufficient flexibility to allow the user to formulate their own applications. Hence the application part of the system is a shell consisting of a number of encapsulated application modules.

3) The integration of hydrologic and agricultural models and a GIS will be the primary activity of this project.

4) The user will interface with a number of data sources. Thus an innovative multi-task user interface based on graphics and visualization is needed.

The proposed project design utilizes public domain software systems which would run with a multitude of software/hardware systems. It also utilizes some of the most comprehensive data bases available to the environmental scientific and commercial community. This, therefore, makes such a system a likely candidate for commercialization.

NASA will team up with the U.S. Army Corps of Engineers and The Houston Advanced Research Center to integrate the models into EIAMS. NASA and HARC scientists have had a productive research association for the past several years.

In the following, Section 2 will discuss the design strategy of EIAMS; Section 3 will detail the critical development areas; and Section 4 will discuss the applications for EIAMS. Sections 5 through 9 will cover the implementation plan, management approach, personnel, budget and available facilities.

2. SYSTEM DESCRIPTION

It is the intent of the proposed project to show the feasibility of using satellite data and associated spatial modeling and analysis system as a unified tool to assist scientists in forecasting potential droughts, impacts on agricultural and ecological systems, and to eventually make recommendations for potentially impacted regions. This requires complex analysis of geographic data which need to be organized to support rapid and effective simultaneous usage.

We have considered integrating a number public domain software systems. For example, GSFC has a number of tools for satellite data ingest and analysis; COE has a GIS and hydrologic models; HARC is providing the agricultural models, and image processing will be done with Khoros, a University of New Mexico public domain image processing package.

The core software packages to be integrated into EIAMS are: 1) software for satellite and environmental data acquisition, 2) software for analyzing and displaying satellite and in situ data (raster and vector on the same display), 3) hydrologic and agriculture models, 4) data management software, 5) geographic data bases and software to integrate them, and 6) a smart user interface. Software systems for 1, 2 and 3 are either existing or are under development at the International Data Systems Office, GSFC or HARC. Hence the critical development areas are in 4, 5 and 6.

EIAMS will be user friendly and flexible to include algorithms and models of specific concern to users (e.g. hydrologic models). It can be configured to run on a single machine or on a distributed network system wherein software modules will be running in different computers and users will have workstations with access to the network.

The modules (or subsystems) in EIAMS and their connectivity can be illustrated in Figure 1. With this configuration, EIAMS receives data from the Global Telecommunication Service and any local environmental and meteorological reports. Its own satellite acquisition equipment is able to receive TIROS and any geostationary satellite (GOES, GMS, METEOSAT and INSAT). The integrated module for satellite and environmental data display can handle both raster and vector graphics at the same time and is also connected to the geographic information module. This analysis and display subsystem includes scientific algorithms such as remote sensing retrievals of atmospheric temperature and water vapor soundings, sea surface temperature, rainfall rate, vegetation index and other geophysical parameters. Initially, the environmental models that are part of EIAMS includes a hydrological model for flood prediction and agricultural models for environmental assessment.

The data flow among the collection of software subsystems is managed by an innovative data management software, which employs a archive manager and reformat compiler to incorporate flexibility into EIAMS. Such flexibility allows the adding or upgrading of geographic information and environmental models. Furthermore, the user will be interacting with a number of software modules. They are, for example, the environmental models, image and data analysis software and a geographic information system. Hence the user interface is designed to allow the user to interact with several processes at the same time.

3. CRITICAL DEVELOPMENT AREAS

A critical component of future Earth science studies will be the development of an interactive spatial information and analysis system. A scientist or a manager/decision maker should not have to learn all of the necessary technologies to process data in order to make informed decisions. No single system is fully capable of performing the wide variety of GIS and spatial processing tasks required for a quick response to environmental events. Therefore, software integration issues are of paramount importance in the selection of components and the design of the overall system. The system should gracefully handle combinations of raster, vector and tabular data while maintaining proper orientation in both space and time.

The wide variety of data sources and software package to be potentially integrated demands careful attention to initial system design, data reformatting programs, and data set format tracking. A flexible data interface to the spatial modeling software will feed the GIS data into the models and feed the model output back into the GIS for generating output products. All data sets must be properly tagged to track time sequences, assessments of the coupling of various parameters, and to compare model output to reality and to those from different or improved models.

Four development areas are considered critical. They are: 1) the development of a data management software, 2) the integration of existing analysis tools, and 3) the integration of models and GIS, and 4) the development of a graphical user-interface These critical development areas will be discussed in the following subsections.

3.1 DATA MANAGEMENT

Several software packages may be relevant in the analysis of an environmental event, each containing features best suited to certain parts of the analysis and each requiring different data and command input. A complete analysis of the environmental event may include communication between processes, for example, in feeding one application package's output to another package as input or combining the results of several models to make predictions. Hence we have to develop an Archive Manager and a Reformat Compiler.

An Archive Manager will serve as a librarian, knowledgeable as to the whereabouts and format of all relevant data. The data can encompass a wide range of media (disk drives, magnetic tape, optical disk, etc.) and cover a wide geographic area, and have varied formats. The Archive Manager will be capable of determining what data is available in response to a simple user query specific to the type of data being searched for. For example, a search could be made for the satellite images covering a particular region, or the rainfall rates during a particular time in a particular region. Thus, queries can be in regard to temporal and spatial as well as to data specific attributes such as pixel resolution.

A potential role of the Archive Manager would be a front end to a relational data base management system (RDBMS). It would communicate between the user and the RDBMS helping to navigate the numerous catalog relations, translating spatial or temporal queries into RDBMS syntax and preserving query contexts and accumulated query results. The Archive Manager would accept varied user query, such as geographic place names instead of geographic coordinate, and translate units such as "feet" or "meters" into appropriate model coordinate value.

A Reformat Compiler produces data in a format suitable for a particular model or analysis module. Responding to a command to transfer data from one package to another, the Reformat Compiler will read a file describing the input format of the data and the desired output format. For many instances, pre-existing data descriptions (known to the Archive Manager) associated with the data sets, and output data formats and associated with the application packages could be used as automatic input to the Reformat Compiler so that data translation could be performed on the fly. The operation of the Reformat Compiler can be summarized by Figure 2.

A comprehensive Reformat Compiler (RC) in needed to reformat and input the numerous varied data sets. Users will be able to add applications to the system. Standard system integration tools such as X-windows, Network File Systems (NFS), and UNIX operating system will be implemented as appropriate. We envision developing a "seamless" software system where various data sets from numerous satellites are being analyzed under one system.

TRENDS IN DATA FORMATS FOR THE SPACE AND EARTH SCIENCES

Dr. Randall Davis
Laboratory for Atmospheric and Space Physics
University of Colorado

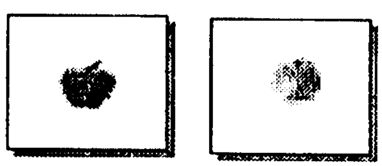
August 11, 1992

Trends in Data Formats for the Space and Earth Sciences

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

NASA Applied Information Systems Research Program
Workshop II
Boulder, Colorado
11 August 1992

What We Have Here Is a Failure to Communicate



- Much of our valuable space and earth science data are stored in ways that severely limit their use by current scientists and their value to future scientists
  - Vital information about the form and content of data files is lacking
  - Similar types of data are implemented in very different ways
  - Computer-dependent formatting (number formats, record formats, etc) and quirky encoding schemes are used with abandon

Changes in Archiving, Distribution and Use of Scientific Data Increase Need for Better Data Formats

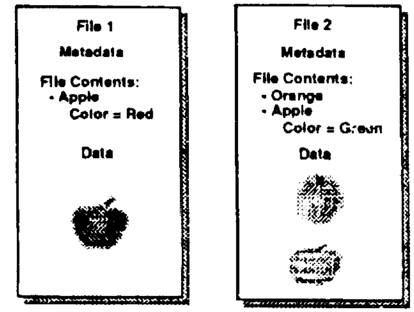
- 1970s — Comprehensive archives
  - World Data Centers
  - ESSA/NOAA, NASA and USGS Data Centers
- 1980s — NASA Discipline Data Systems
  - 2a: Pilot systems (PCDS, PLDS, PODS and PPDS)
  - 2b: Operational systems (ADS, NCDS, NODS and PDS)
- 1990s — Consolidation of NASA's earth-oriented discipline data systems (NCDS, NODS, and PLDS) into the Earth Observing System Data and Information System (EOSDIS)
- 2000 — Increased uniformity between major U. S. earth science data systems (NASA, NOAA and USGS) and across nations

NASA held an invitational data format workshop in June to begin to determine if modern data formats will meet the needs of the future

Formats Discussed at the Workshop

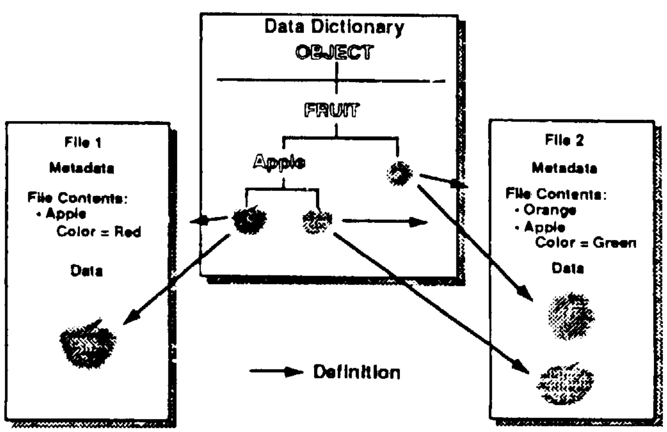
- GRIB and BUFR
  - World Meteorological Organization
- Common Data Format (CDF)
  - National Space Science Data Center
- NetCDF
  - UCAR Unidata
- Flexible Image Transport Standard (FITS)
  - International Astronomical Union
- Hierarchical Data Format (HDF)
  - National Center for Supercomputing Applications
- Spatial Data Transfer Standard (SDTS)
  - U. S. Geological Survey
- CEOS Superstructure
  - Committee on Earth Observing Satellites
- Standard Format Data Unit (SFDU)
  - Consultative Commitee for Space Data Systems
- Planetary Data System (PDS)

The Common Thread: Formats Store Identifying and Descriptive Metadata

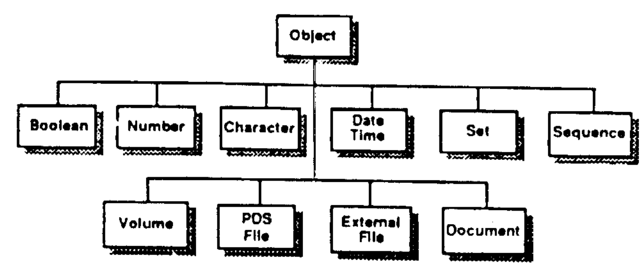


- Identification and location of data objects within a message, file or volume can be specified with tags — machine-friendly encoded descriptors — or labels in a human and machine-readable language
- *Attributes* describe important characteristics of data objects; they are often formally required by documentation or defined by a data dictionary

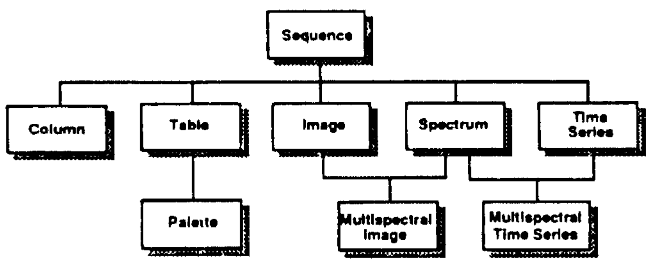
PDS and SFDU Approaches Promote Re-Usable Formal Data Types



Examples of Formal Data Types: A Partial View of the PDS Data Object Hierarchy



More PDS Objects: A Partial View of the Sequence Classes



Conclusions, Results and Action Items From the Data Formats Workshop

- Conclusions:
  - Good data formats are available for space and earth science applications
    - More and more projects will require data in specific formats (for example, EOSDIS Version 0 is currently planning on HDF data format with metadata in PDS label format)
  - The relationship between special scientific data formats and data formats from general computing has to be examined further
- Results:
  - Developers of CDF, netCDF and HDF are examining the possibility of developing a common interface to data in their formats
  - Future workshops will likely be held
- Action Items:
  - Clarify the kinds of objects that are supported by each format
  - Determine the utility of translators between specific formats
  - Foster greater pan-discipline cooperation in developing or selecting formats

THE GRID ANALYSIS AND DISPLAY SYSTEM (GrADS)

Dr. James L. Kinter III
Center for Ocean-Land-Atmosphere (COLA) Interactions
University of Maryland

August 12, 1992

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) 405-5389 (5395)
doty@cola.umd.edu
kinter@cola.umd.edu

CENTER FOR OCEAN-LAND-ATMOSPHERE INTERACTIONS

Basic Research in:

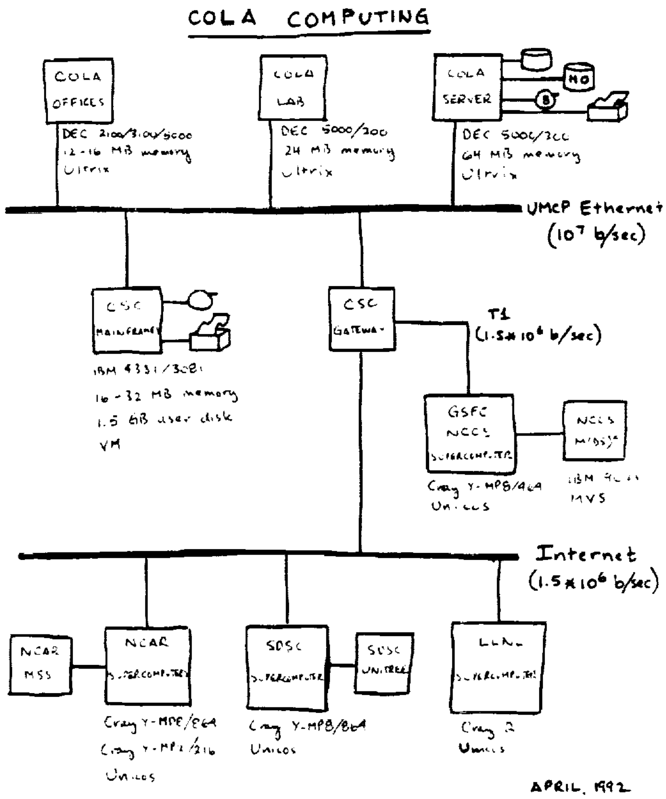
- Climate Modeling
  - Monthly to Seasonal Predictability
  - Interannual Variability
  - Deforestation
  - Desertification
  - Satellite Sensor Accuracy Requirements
- Data Assimilation
  - Reanalysis for Climate Diagnostics/Monitoring
- Numerical Weather Prediction
  - Dynamic Extended Range Forecasting (DERF)
- Ocean Modeling
  - Tropical Atlantic Variability
  - Coupled Climate Modeling
  - Use of New Sensors
- Dynamical Systems
  - Predictability Theory
- Information Systems
  - Visualization, Data Manipulation and Management

CENTER FOR OCEAN-LAND-ATMOSPHERE INTERACTIONS

- Atmospheric General Circulation Model (AGCM)
  - R40L18 resolution with full physics and diurnal cycle
  - Simple Biosphere Model
  - Each seasonal simulation
    - requires 25 hours Cray Y-MP time
    - requires up to 64 MB memory on one processor
    - generates 2.7 GB data
- Oceanic General Circulation Model (OGCM)
  - Based on GFDL Modular Ocean Model
  - Each seasonal global simulation
    - requires 5 hours Cray Y-MP time
    - requires up to 32 MB memory on one processor
    - requires 120 MB SSD (Cray Solid State Disk)
    - generates 1 GB data
- Global Data Assimilation System
  - Based on NMC GDAS
  - Each seasonal reanalysis
    - requires 140 hours Cray Y-MP time
    - requires up to 64 MB memory on one processor
    - generates 4 GB data

===> SUBSTANTIAL SUPERCOMPUTER TIME REQUIRED

===> DATA MANAGEMENT & VISUALIZATION CRITICAL



Complaints About Existing Scientific Visualization Tools

(adapted from results of a survey conducted by Mike Botts)

* Tool is not extensible or flexible
* Tool is too difficult to learn and use
* Too difficult to get existing data into the tool
* Tool does not adequately link processes of visualization and analysis
* Problems with heterogeneous environments
* Tool designed with computer graphics, rather than science, in mind
* Scientists not aware that the tool exists
* Scientists lack appreciation that the visualization tool is a serious scientific tool
* Too difficult to communicate the results

GrADS - DESIGN GOALS

INTEGRATE:

ACCESS	MANIPULATION	DISPLAY
4D Gridded Data Station Data General Slices	Expressions Functions	Maps Charts Animation

INTERACTIVE:

Sub-Second Response
Data, Display Control
Scripting, Programmability

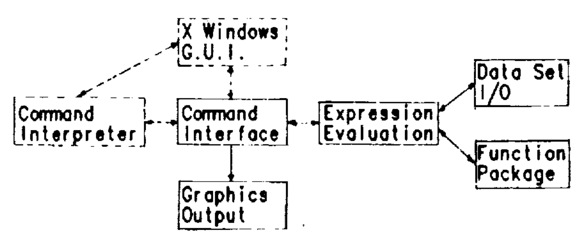
EASE OF USE:

Easy to Learn
Intuitive

HARDCOPY:

Vector Graphics

Design



Page 2 COLA/UMCP

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 - IMPLEMENTATION

PORTABILITY:

ANSII Standard C Code
Isolated Device-Specific Graphics
Standard Unix "Stream" Datasets

HIGH SPEED GRAPHICS:

Built-In Routines
Optimization

EXPANDABLE FUNCTIONALITY:

Easy to Add Graphical Displays
Easy to Add Data Manipulation Functions

GrADS - REQUIREMENTS

HARDWARE:

Moderately Fast Processor
(386/7, any RISC)

Moderate Resolution Display
(VGA or better)

Dynamic Memory Usage (0.5 MB Minimum)

3 MB disk for GRADS executables and map files

Enough Disk Space for User Data
(Local or on Network Server)

Postscript Printer - Color or Monochrome
(other devices can be easily supported)

SOFTWARE:

ANSII Standard C Compiler + Libraries

X Windows (UNIX)

GrADS DATA SETS

- GrADS INTERNAL FORMAT
  - Binary
  - Optimized for I/O Performance
- 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.)

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

GrADS - CURRENT USER GROUPS

(in chronological order of contact)

U.S. UNIVERSITIES AND FEDERAL LABORATORIES

University of Maryland at College Park
Columbia University (Lamont-Doherty Geological Observatory)
Colorado State University
California Institute of Technology (NASA JPL)
National Center for Atmosphere Research
UNIDATA (University Corporation for Atmospheric Research)

NASA Goddard Space Flight Center (Codes 910.3, 910.4, 913)
NOAA/National Meteorological Center (Dev. Div. and CAC)
NOAA/ERL Geophysical Fluid Dynamics Laboratory
NOAA/ERL Air Resources Laboratory
NOAA/ERL Forecast Systems Laboratory
U.S. Geological Survey (USGS - Reston, VA)

U.S. COMMERCIAL INSTITUTIONS

Atmospheric and Environmental Research, Inc.

FOREIGN UNIVERSITIES AND GOVT. LABORATORIES

INPE/CPTEC (Space Studies Institute, Brazil)
CNR/IMGA (Geophysics Institute, Italy)
ICTP/ICS (Climate Institute, Italy)
ENEA (Energy Agency, Italy)
National Tidal Facility (Australia)
KIST/SERI (Korea)
Dalhousie University (Canada)

NATIONAL CENTER FOR ATMOSPHERIC RESEARCH (NCAR) INTERACTIVE STATUS

Dr. Bob Lackman
NCAR

August 12, 1992

NASA/AISRP Workshop
August 11-13, 1992
Boulder, Colorado

NCAR Interactive Status

by

Bob Lackman

SCIENTIFIC VISUALIZATION GROUP, SCD
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

Collaboration Review

- NCAR Interactive - NCAR/SCD
- ENVISION - U of Illinois
- GrADS - U of Maryland
- PolyPaint - NCAR/MMM
- netCDF - UCAR/Unidata

SCIENTIFIC VISUALIZATION GROUP, SCD
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

NCAR Interactive Community Goals

- Guarantee long term support via cost recovery
- Provide university and non-profit researchers low cost visualization
- Advance and support the scientific infrastructure through common software

SCIENTIFIC VISUALIZATION GROUP, SCD
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

NCAR Interactive Design Goals

- Build on existing NCAR Graphics libraries
- Maintain a single package for distribution
- Retain portability across systems

SCIENTIFIC VISUALIZATION GROUP, SCD
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

NCAR Interactive Design Goals (cont.)

- Three interface levels

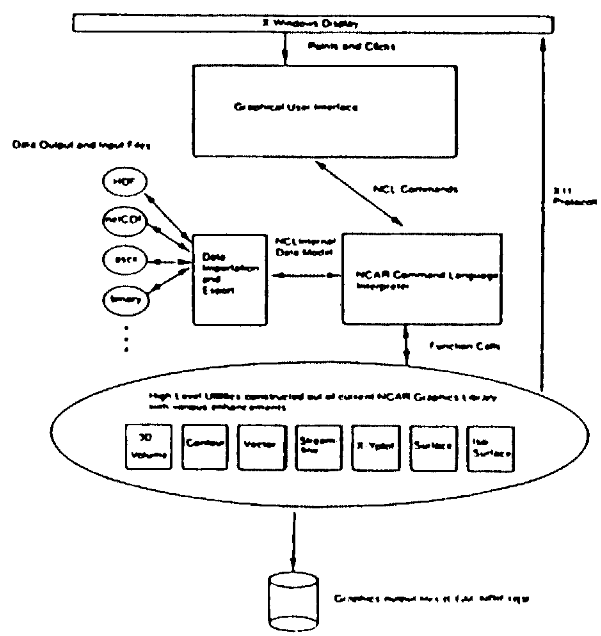
  Programmatic
    High Level Utilities (HLUs)
    C and Fortran bindings

  Command line
    NCAR Command Language (NCL)

  Visual point and click
    Graphical User Interface (GUI)

SCIENTIFIC VISUALIZATION GROUP, SCD
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

FIGURE 1 System Structure



NCAR INTERACTIVE Functional Specification

Project Schedule

- 1st year summary, August 1992

  Functional requirements
  Preliminary design
  Prototype NCL for netCDF data access
  A prototype HLU

SCIENTIFIC VISUALIZATION GROUP, SCD
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

Project Schedule (cont.)

- 2nd year objective

NCL Command Interpreter
Multiple HLUs
Prototype GUI

SCIENTIFIC VISUALIZATION GROUP, SCD
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

Project Schedule (cont.)

- 3rd year objective

Distribution in Version 4.0
of a limited functionality
working system

SCIENTIFIC VISUALIZATION GROUP, SCD
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

Using NCAR Graphics in a Data Flow Environment

SCIENTIFIC COMPUTING DIVISION
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

NCAR Interactive Functionality

Ethan Alpert and Jeff Boote
SVG

SCIENTIFIC COMPUTING DIVISION
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

1

Project history

- Functional requirements from Aug. 91 - present
  - Prototype development and demonstration
  - Functional specification document
  - Personal interviews with NCAR scientists
- Design from Apr. 92 - present
  - Outlining changes needed for current utilities

SCIENTIFIC COMPUTING DIVISION
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

2

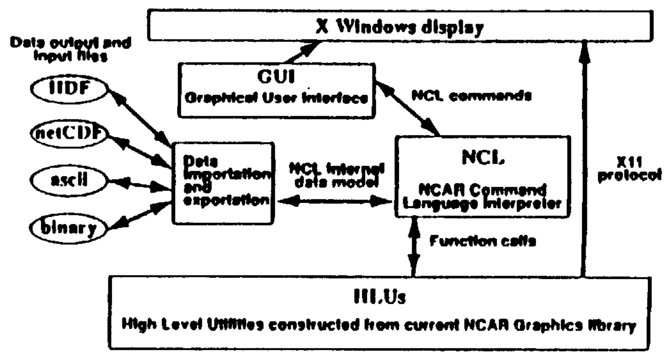
What do we hope to accomplish?

- Support users with multiple skill levels
- Decrease time spent specifying visualizations
- Create an infrastructure that we can build on
- Support interactive data analysis and exploration
- Build a system that will handle data in multiple formats
- Provide mechanisms for users' contributions

SCIENTIFIC COMPUTING DIVISION
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

3

What will NCAR Interactive look like?



SCIENTIFIC COMPUTING DIVISION
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

4

What is NCL?

- NCAR Command Language
- An interpreted command language
- Uses some language constructs from FORTRAN 90
- Contains special language constructs specific only to NCL
- Supports multiple data formats for input

SCIENTIFIC COMPUTING DIVISION
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

5

What is NCL?

- Variables - scalar and array, located in files or in memory

a = 1.0
b = [1.0, 2.0, 3.0; 4.0, 5.0, 6.0; 7.0, 8.0, 9.0]
c = b(1:2,2:3)

2.0 3.0
5.0 6.0

d = Temp691.temp(1:51, 1:100:4)

SCIENTIFIC COMPUTING DIVISION
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

6

What is NCL?

- Built-in algebraic operators: * / ^ # + - < >

a = [2,2;3,3]
b = [4,5;6,7]
c = a * b

8 10
18 21

SCIENTIFIC COMPUTING DIVISION
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

7

What is NCL?

- Loops - while and do

```
while ( a LT 10 )
        a = a + 1
endwhile

do a = 1, 10, 2
        x = x + a ^ 2
endo
```

SCIENTIFIC COMPUTING DIVISION
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

8

What is NCL?

- If statements

```
if ((a EQ b) and (c LT d))
        a = c
else
        a = b
endif
```

SCIENTIFIC COMPUTING DIVISION
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

9

What is NCL?

- Functions and procedures

```
addfile("/u1/ethan/data/Temp691.cdf")
a = sqrt(Temp691.temp)
```

SCIENTIFIC COMPUTING DIVISION
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

10

What is NCL?

- Functions and procedures can be:
    - User-defined in NCL source
    - Built-in math and system functions
    - Built from user FORTRAN or C source
- NCL operates in both line-by-line immediate mode and a batch script mode

SCIENTIFIC COMPUTING DIVISION
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

11

What is unique about NCL?

- Data model similar to netCDF data model
  - Files contain one or more variables
  - Files contain named dimensions used to define shape of variables
  - Variables can contain descriptive attributes
  - Each dimension can have an associated coordinate variable

SCIENTIFIC COMPUTING DIVISION
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

12

What is unique about NCL?

- Special language constructs that support the data model
  - Coordinate indexing

```
a = Temp691.temp({10:60}, {-50:-150})
```

@ & !

```
a@units = "Degrees C"
a&lat = (90,85, 80, . . . )
a!0 = "lat"
```

SCIENTIFIC COMPUTING DIVISION
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

13

What is unique about NCL?

- Random access to data on disk

```
a = Temp691.temp({10:60},{-50:-150})

temp = Temp691.temp
a = temp({10:60},{-50:-150})

Temp691.temp({10:60},{-50:-150}) = a({10:60},{-50:-150})
```

SCIENTIFIC COMPUTING DIVISION
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

14

What is unique about NCL?

- Support for format-independent I/O
  - Requirements for adding a format
    1) Format supports random access read/write
    2) A convention can be established for handling coordinate variables, named dimensions, and attributes

SCIENTIFIC COMPUTING DIVISION
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

15

What is unique about NCL?

- Visualization specification block

```
visblk ncar mycontour {
        NgbPlotStyle: CONTOUR
        NmpMap: True
        NmpOutline: PS
        NmpProjection: OR
        NvpX: .1
        NvpY: .1
        NvpWidth: .8
        NvpHeight: .8
}
assigndata(mycontour, Temp691.temp)
```

SCIENTIFIC COMPUTING DIVISION
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

16

What is unique about NCL?

- Attribute sets for function invocation

```
function ctof(x:float:{units = "Degrees C"})
begin
        . . .
end
```

SCIENTIFIC COMPUTING DIVISION
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

17

What is unique about NCL?

- Support for propagation of missing values

SCIENTIFIC COMPUTING DIVISION
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

18

NCL summary

- NCL provides interactive data access for NCAR Interactive
- NCL provides an extensible environment for processing data prior to visualization
- NCL will allow sites to develop custom data storage conventions and still use NCAR Interactive

SCIENTIFIC COMPUTING DIVISION
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

19

NCAR Interactive's

Graphical User Interface

A plot tool

SCIENTIFIC COMPUTING DIVISION
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

20

What will the GUI do for you?

- Give nonprogrammers access to NCAR Graphics
- A WYSIWYG tool for producing publication-quality graphics
- Allow interactive data exploration directly from plots

SCIENTIFIC COMPUTING DIVISION
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

21

Basic use of GUI

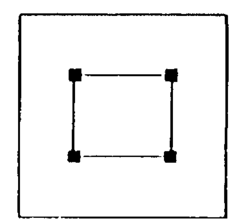
1. Creating plots
2. Changing characteristics of plots
3. Accessing data
4. Associating data with a plot
5. Data exploration

SCIENTIFIC COMPUTING DIVISION
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

22

Creating plots

1. Choose a plot style
2. Move the bounding box on the frame to the position desired



SCIENTIFIC COMPUTING DIVISION
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

23

Changing plot characteristics

1. Resize and move directly on frame
2. Plot specification window
   - Common characteristics - maps & tick marks
   - Specific characteristics

SCIENTIFIC COMPUTING DIVISION
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

24

Accessing data

- Defining data cells
  1. Specify data file or other cell as input
  2. Specify input types
  3. Specify output types
  4. Write NCL script to process data

SCIENTIFIC COMPUTING DIVISION
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

25

Associating data with a plot

1. Select a cell for the plot
   - Only cells that have the correct output types will be allowed
2. Indicate which dimension of the data will be tied to each dimension of the plot

SCIENTIFIC COMPUTING DIVISION
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

26

Data Exploration

- Direct selection on plot
- Edit data values - missing values

SCIENTIFIC COMPUTING DIVISION
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

27

Animations

- Record feature
- Playback
- Speed

SCIENTIFIC COMPUTING DIVISION
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

28

Colormap editor

- Create/edit entire colormaps
- Set of default colormaps
- Edit individual cells of a colormap
- Indicate the plots that are using specific cells

SCIENTIFIC COMPUTING DIVISION
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

29

GUI configuration

- Allow dynamic changing of most "X Resources"
- Allow saving of a given configuration so future invocations of the GUI automatically start the same way

SCIENTIFIC COMPUTING DIVISION
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

30

Help interface

- Contextual help
- Index of help

SCIENTIFIC COMPUTING DIVISION
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

31

Saving state

- Allow the user to exit the program and continue at the same place at a later time

SCIENTIFIC COMPUTING DIVISION
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

32

Creating default plot styles

- Allow the user to simply create textual plot default files
- Files can be used at almost any level of NCARI

SCIENTIFIC COMPUTING DIVISION
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

33

High Level Utilities (HLUs)

- The HLUs will be a new programming interface to NCAR Graphics
- The HLUs will provide consistency between utilities
- The HLUs will provide a visualization model to the user, which will simplify the use of NCAR Graphics as a programming toolkit
- The HLUs are an alternative to the current Fortran interface, not a replacement

SCIENTIFIC COMPUTING DIVISION
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

34

The HLU visualization model

- Steps for creating and specifying a plot with the HLUs

1. Create an instance of the style of plot you want
   This returns a unique ID that identifies the plot
2. Configure plot - defaults file or setarg
3. Draw plot
4. Destroy plot or modify configuration and update

SCIENTIFIC COMPUTING DIVISION
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

35

The HLU visualization model

- What are the differences between the HLUs and current NCAR Graphics?
  1. Users no longer have to be familiar with the ordering of function calls to produce output
  2. Users no longer have to be familiar with GKS
  3. Details of plotting hidden from users
  4. Consistency

SCIENTIFIC COMPUTING DIVISION
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

36

The HLU visualization model

- Another feature of the HLUs is to integrate related utilities that are often used together

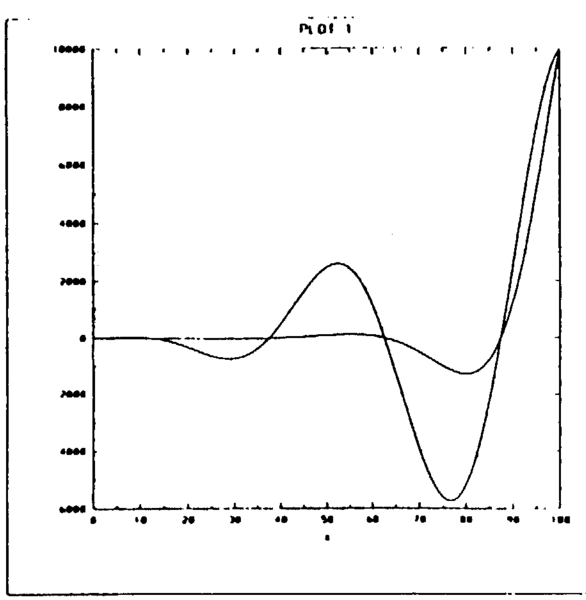
  For example: CONPACK, EZMAP, PLOTCHAR, LABELBAR, DASHCHAR, SPPS, and GRIDAL must often be used together to make one contour plot

- Similar functionality is presented in an identical fashion in different utilities

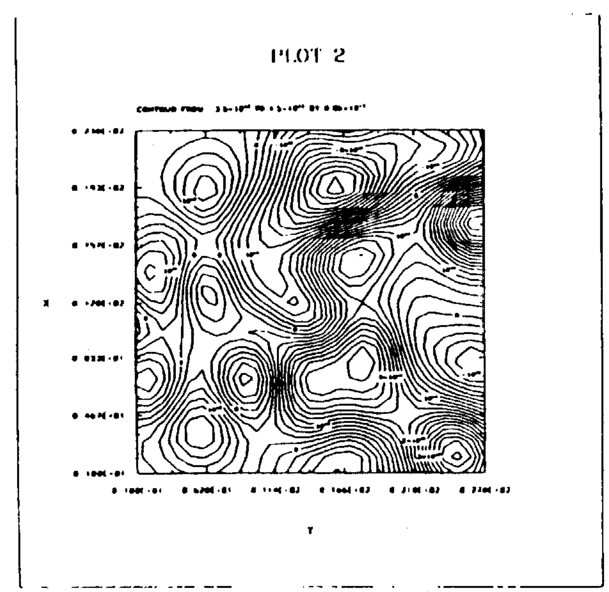
  Consider similarities between CONPACK plot and AUTOGRAPH plot

SCIENTIFIC COMPUTING DIVISION
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

37

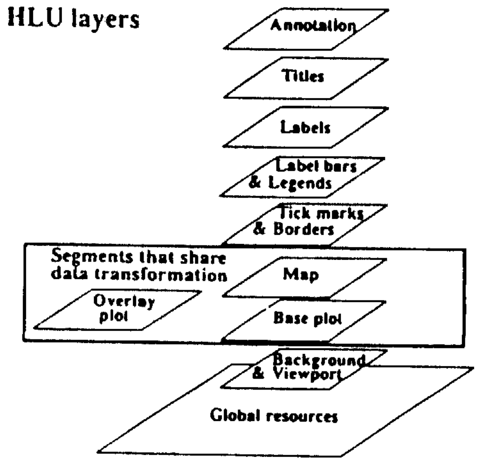


38



39

HLU layers



SCIENTIFIC COMPUTING DIVISION
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

40

The HLU visualization model

- Resources used to configure options for different layers

  NtmXStart - Sets value to begin tick marks on
  NtiMainText - Sets string for main plot title
  Nmp - Turns on map drawing

- Resources set with setarg and getarg functions
- Every 1D and 2D HLU has same names for common resources

SCIENTIFIC COMPUTING DIVISION
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

41

The HLU visualization model

- Draw command uses all the resources set for a given instance and renders the plot
- If resources are not set, system defaults are used
- Some resource values are generated automatically

  For example: a color-filled contour plot will automatically generate a label bar with the appropriate color and range entries

SCIENTIFIC COMPUTING DIVISION
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

42

The HLU visualization model

- Naming convention for common resources

Prefix	Resource class
Nan	Annotation
Nti	Title
Ntm	Tick mark
Nlb	Label bar
Nmp	Map
Nbk	Background
Nvp	Viewport
Nlg	Legend
Nbo	Border
Ngb	Global

SCIENTIFIC COMPUTING DIVISION
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

43

The HLU visualization model

- Naming convention for specific HLU plot types

Prefix	HLU name
Ncn	Contour
Nvr	Vector
Nsr	Surface
Nis	Isosurface
Nsl	Streamline
Npp	PolyPaint
Nxy	X-Yplot
Nhs	Histogram
N3d	Common Surface and Isosurface

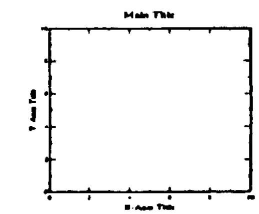
SCIENTIFIC COMPUTING DIVISION
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

44

The HLU visualization model

- User defaults file

Nvpx	10
Nvpy	90
NvpWidth	80
NvpHeight	80
NtiMain	Main Title
NtiMainFont	Times-Roman
NtiMainFontSize	04
NtiX	X-Axis Title
NtiXFont	Times-Roman
NtiY	Y-Axis Title
NtiYFont	Times-Roman
NtiXFontSize	03
NtiYFontSize	03
NtmXGroup	True
NtmYGroup	True
NtmXMajorLength	009
NtmXMinorLength	0045
NtmXMinorPerMajor	1
NtmXLabelFont	Times
NtmXLabelFontSize	02
NtmYMajorLength	008
NtmYMinorLength	0045
NtmYMinorPerMajor	1



SCIENTIFIC COMPUTING DIVISION
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

45

The HLU visualization model.

- Overlays
- Plots to be overlaid are registered with a main master plot
- Master plot then configures overlay plots' viewport and data transformation resources
- After master plot is drawn, overlays are drawn on top of the master plot in the order they were registered

SCIENTIFIC COMPUTING DIVISION
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

46

The HLU visualization model

- Annotation

  Arrows, Lines, Circles, Boxes, Text, Legends, Label bars, and more

- Annotations are registered with an instance of a plot and drawn when the draw function is called

SCIENTIFIC COMPUTING DIVISION
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

47

HLU Summary

- HLUs are a new programming interface
- HLUs are resource-driven (argument-driven)
- There is a definite visualization model presented
- The HLUs will not replace the existing utilities

SCIENTIFIC COMPUTING DIVISION
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

48

Summary

- Support users with multiple skill levels
- Decrease time spent specifying visualizations
- Create an infrastructure that we can build on
- Support interactive data analysis and exploration
- Build a system that will handle data in multiple formats
- Provide mechanisms for users' contributions

SCIENTIFIC COMPUTING DIVISION
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

49

It is your package too!

- Evaluate prototypes
- Notify us of your changing needs!

Ethan Alpert ethan@ncar.ucar.edu

Jeff Boote boote@ncar.ucar.edu

SCIENTIFIC COMPUTING DIVISION
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

50

Using NCAR Graphics in a Data Flow Environment

I. Introduction

II. Issues addressed:

a. Data Flow Environment Definition and examples
b. Pros and Cons of Using NCAR Graphics in a Data Flow Environment
c. Future enhancements to NCAR Graphics for Data Flow Environment support
d. Hints on importing NCAR Graphics into a data flow environment

III. Definition of a data flow Environment?

IV. Module Types

a. Read - used to read your data format (netCDF, ascii, ...)
b. Filter - used to subsample, massage and process your data
c. Map - used to convert data into a visual representation
d. Display - used to render your image or output.

V. Popular Examples of Data Flow Environments:

a. apE -(animation production Environment) originally developed at the Ohio Super Computer Center, Now TaraVisual Corp.
b. Khoros - University of New Mexico
d. Explorer - Silicon Graphics
e. AVS (Application Visualization System) - Advanced Visual Systems. Inc.

VI. Benefits of Using NCAR Graphics in a Data Flow Environment

a. Ease of use for "Non-programmers"
b. Widget Customization
c. Distributed Processing.
e. Image Processing Tools
f. Public Domain Modules
   i. International AVS Center (avs.ncsc.org) N. Carolina Super Computing Center
   ii. UK and US site for Explorer modules swedishchef.lerc.nasa.gov
g. NCAR Graphics adds more power and flexibility to Data Flow Environments

VII. Discussion of reasons for not using NCAR Graphics with a Data Flow Env.

a. NCAR Interactive
b. NCAR Graphics functionality already in Data Flow Environments
c. Cost
d NCAR Graphics does not currently support geometric data output.

VIII. Importing NCAR Graphics into a Data Flow Environment Module

a. Module Components:
   i. Parameter input specification (widgets)
   ii. Data input specification (input port type)
   iii. Data output specification (output port type)
   iv. Computation routine (NCAR Graphics program)
b. Module Building Templates

IX. Displaying NCAR Graphics output in a Data Flow Environment

X. Summary

A system with an advanced graphical user interface where programs are built by connecting modules in a pipeline

Data Flow

Popular Data Flow Environments

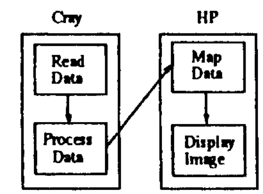
- apE (Application Production Environment) from Tara Visual Corporation
- AVS (Application Visualization System) from Advanced Visual Systems, Inc.
- Explorer from Silicon Graphics, Inc.
- Khoros from the University of New Mexico

What are the benefits of using NCAR Graphics in a Data Flow Environment?

- "Non-programmers" can quickly and easily create complex NCAR Graphics applications without writing any code

- It is easy to customize a widget to control parameters to your NCAR Graphics routines.

- Distributed Processing



- You can label your NCAR Graphics images with a point and click interface.

- You can easily combine, pan, and zoom NCAR Graphics images.

- You can access a growing database of public domain modules.

Anonymous ftp Sites

AVS
avs.ncsc.org

Explorer
swedishchef.lerc.nasa.gov

- NCAR Graphics adds more power and flexibility to data flow environments.

What are the disadvantages of using NCAR Graphics in a data flow environment?

- NCAR Interactive will provide many of the same capabilities as a data flow environment.

- Data flow environments already have NCAR Graphics-like functionality.

- Cost

- NCAR Graphics does not currently support geometric data.

Steps for Importing NCAR Graphics into a Data Flow Environment

<u>Module Components</u>

1. Parameter Input Specification
2. Data Input Specification
3. Data Output Specification
4. Computation Routine

How to display NCAR Graphics output in a data flow environment

A DISTRIBUTED SYSTEM FOR THE VISUALIZATION AND ANALYSIS OF OBSERVED AND MODELED METEOROLOGICAL DATA

Dr. Steven Koch
NASA/Goddard Space Flight Center

August 12, 1992

A Distributed System for the Visualization and Analysis of Observed and Modeled Meteorological Data

---

A joint effort of
NASA / GSFC (Goddard Space Flight Center) and NCSA (National Center for Supercomputing Applications)

Steven Koch, GSFC
Robert Wilhelmson, NCSA

John Hagedorn, GSFC
Matthew Arrott, NCSA
Gautam Mehrotra, NCSA
Crystal Shaw, NCSA
Jeffrey Thingvold, NCSA

National Aeronautics and
Space Administration

NASA

Goddard Space Flight Center
Greenbelt, Maryland
20771

Telephone : (301) 286-7188
FAX : (301) 286-4661
E-mail : koch@clena.gsfc.nasa.gov

Reply to Attn of: Mail Code 912

Enclosed is a copy of our presentation material for the Advanced Information Systems Research Program Workshop held in Boulder, Colorado on Wednesday, August 12, 1992 titled "A Distributed Analysis System for Model and Observational Data".

Steven E. Koch
Research Meteorologist

Goals

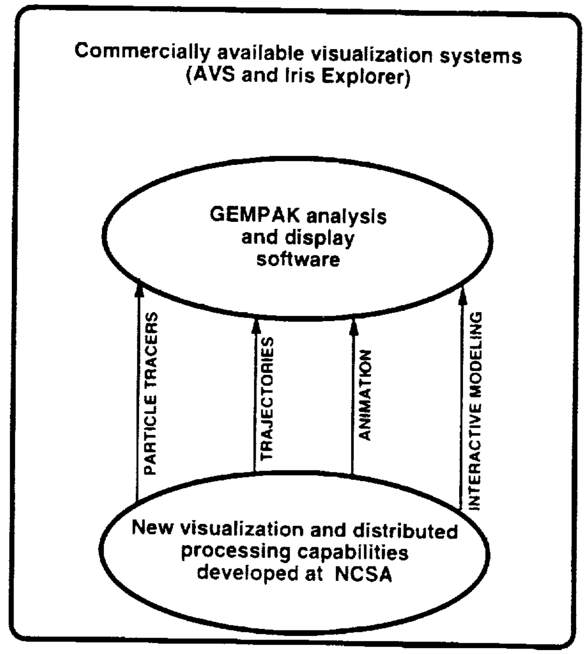
- To create a tool for handling the large amounts of data generated by satellites, observational field programs, and model simulations
- To extend existing 2D mapping capabilities with new analysis functions and modern techniques of 3D visualization, user interaction, and animation.

Approach

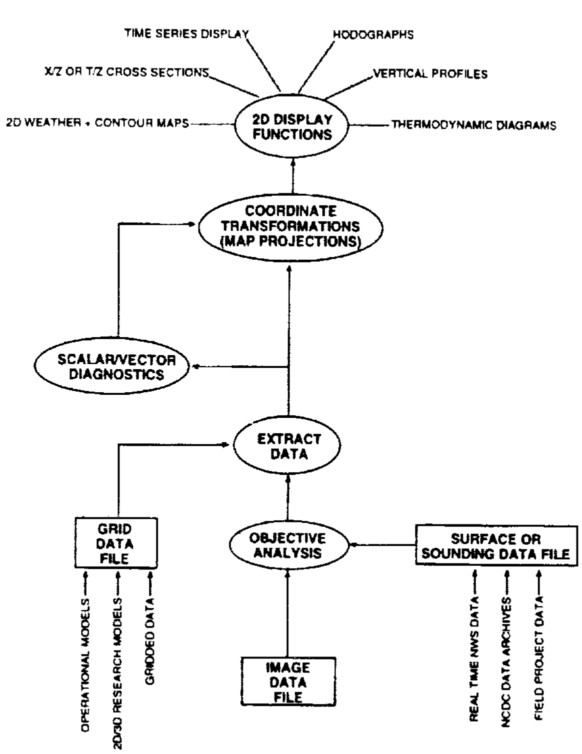
- Maximize the use of existing software
  - GEMPAK, meteorological analysis and display software developed at NASA / GSFC
  - Various 3D interactive capabilities developed at NCSA
- Use commercially available visualization and application builder tools (IRIS Explorer and AVS)
- Provide visualization and analysis capabilities in these areas:
  - 3D volumes of data (isosurfaces, 3D cross sections, vector displays, etc.)
  - Evolution of data over time (time sequences, trajectories, animation, etc.)
  - Distributed processing over workstations and supercomputers, including interactive control of simulations
- Provide a highly interactive environment on a single display

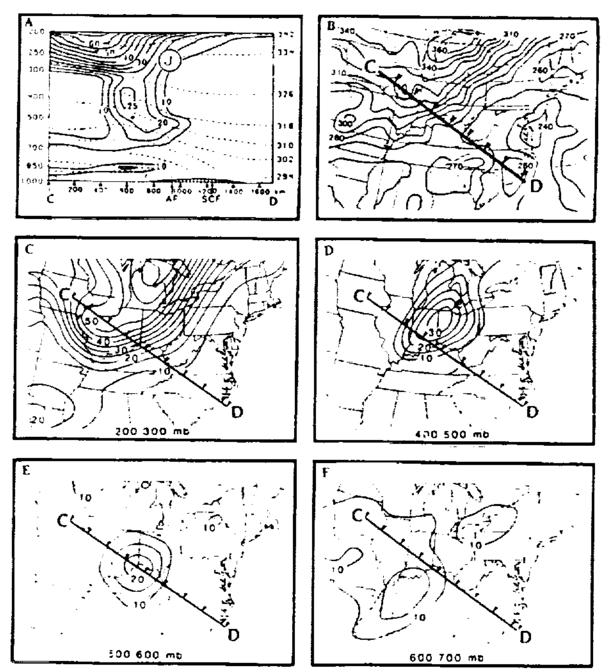
GEMVIS: A Distributed System for the Visualization and Analysis of Observed and Modeled Meteorological Data

A joint effort between NASA/GSFC and the National Center for Supercomputer Applications funded by OSSA/Applied Information Systems Program



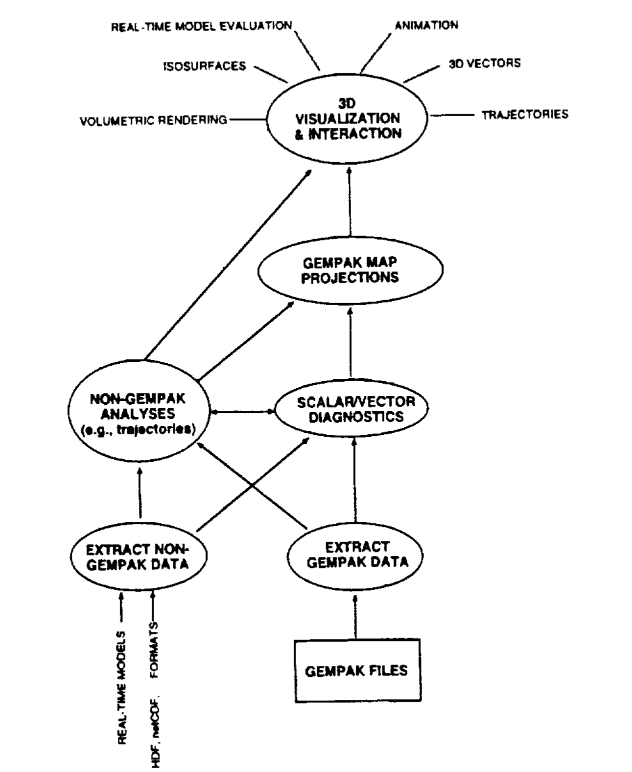
Present 2D GEMPAK Functions

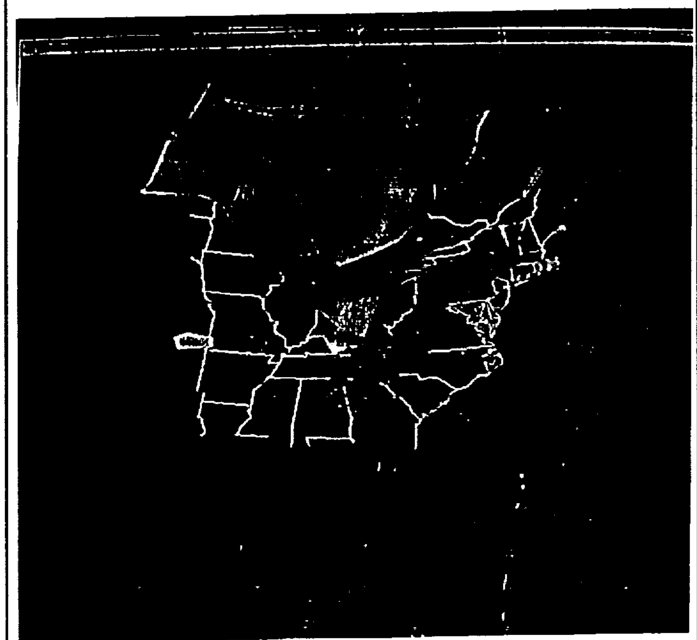




Two-dimensional depictions of a tropopause fold (potential vorticity) using GEMPAK vertical cross section and horizontal contour mapping functions

Planned GEMPAK Functions within the 3D Visualization System





Software Environment

Goal:

Create a flexible, modular, and distributed environment for data handling, model simulations, data analysis, and presentation to be used in studying atmospheric flows on all scales.

Requirements:

A software technology framework which incorporates the following as fundamental design characteristics:

- Change (extensibility, user requirements, functionality, hardware, etc.)
- Integration with Other Solutions (hardware/software systems, output media, etc.)
- Tiered Access for Different Users
- Tools for Constructin User Interface
- Distributed Computing
- Portability
- 3D Interactive Performance

Software Environment (cont.)

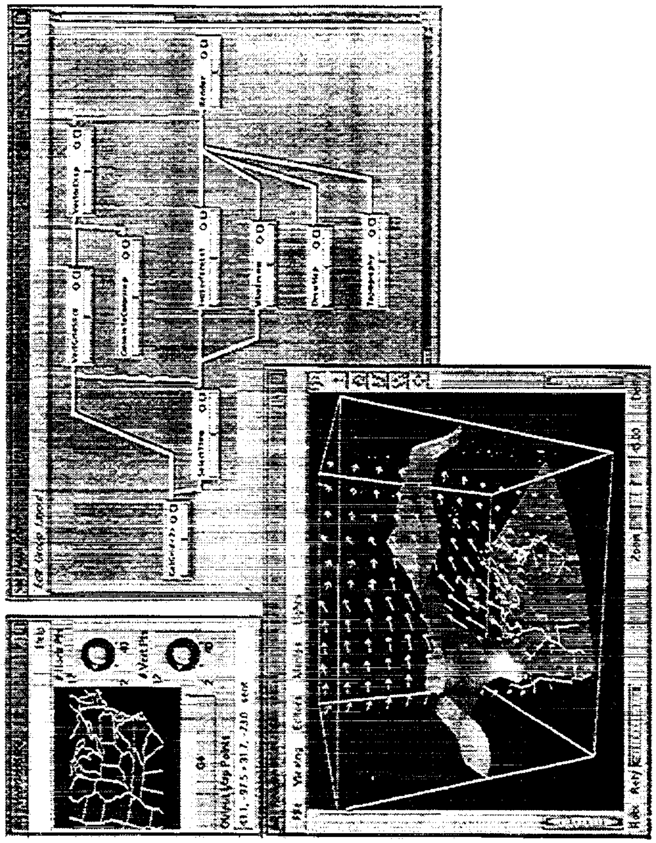
Solution:

Likely environments:
- AVS
- Khoros
- IBM's Data Explorer
- SGI's Explorer

None is proven. Of these, AVS is the most widely used and NCSA has been working with AVS for 2 years. The others are promising - need investigation.

We choose SGI Explorer as our primary tool for these reasons:
- 2nd Generation Design
- Natural distributed processing with natural concurrent execution framework
- Application User Interface designer
- Easily Extensible
- Impressive Development Team
- NCSA in a position of influence as to the direction of Explorer's development

GSFC and NCSA spent time up front to prove feasibility of translation between Explorer and AVS modules, to safe-guard the decision to use Explorer over AVS.



Accomplishments (GSFC)

- Prototype GEMPAK adaptation to AVS, including port of GEMPAK to Convex C220
- Evaluated and selected visualization and application builder software
- Adaptation of GEMPAK to Explorer
  - GEMPAK grid file interface to Explorer
  - Grid scalar and vector diagnostics
  - Coordinate transformations / map projections
  - Topography and map displays
- Modification of GEMPAK to generate 3D vector data
- Display of 3D vector fields

Accomplishments (NCSA)

- Tested alpha and beta releases of Explorer
- Prototype 3D interactive display of multiple level horizontal 2D contours over a map
- Interactive arbitrary vertical cross-section using map locator selection
- Prototype image flip-book animation
- Particle trajectory calculation and display
- Time handling within Explorer, including time-based interpolation and looping constructs
- HDF data file interface to Explorer
- Real time control of simulation running on Cray from within Explorer on workstation
- Demonstration at SIGGRAPH '92

Problems Encountered

- GEMPAK was conceived as an integral system. It's hard to separate subsystems due to unexpected and hard to track connections. These problems have been resolved.
- Explorer Version 1.0 had problems, for example:
  - Bugs in:
    - module grouping
    - synchronization of data from various sources
  - Lacks functionality in:
    - database management
    - animation (scripting, looping, etc.)
    - user interface management
    - annotation

  Explorer Version 2.0 has addressed all of these problems except for database management, which we are addressing in our development.
- Explorer has no support for "missing" data. We will develop our own tools for handing this.
- "Visual Programming Environment" sounds good, but can be confusing to the user. We are using the Explorer module grouping tool to manage the complexity of the visual programs.

What We've Gained Over 2D GEMPAK

- 3-Dimensional Views
- High Level of Interactivity
- Distributed Processing
- Easy Extensibility

What We've Lost From 2D GEMPAK

- Annotation and Physical Coordinate Display
- Standard Meteorological Symbols
- Ability to Run on Low-Performance Machines

Future Work - Near Term

We will release software for Use in Severe Storms Branch of NASA / GSFC in October, 1992.

- Minor User Interface Improvements and Fixes
  - creation of a set of "canned" processing maps
  - module grouping to simplify the interface
  - minor changes to improve consistency among module interfaces
- Preliminary Annotation
  - axis labels
  - titles
  - limited control of size, placement, etc.
- Image Loop Animation
  - the ability to store a series of rendered images either in memory or on disk
  - the ability to "play back" the sequence of images with some control over speed and order of images
- User Documentation

Future Work - Long Term

- Database for Meta-Data
  - units of data
  - time tag
  - physical coodinates
  - etc.
- User Interface Enhancements
  For example, self-configuring control panels
- Enable the User to Query Quantitative Data from the 3D Scene
- Additional Visualization Techniques
  For example:
  - better use of transparency
  - different vector representation
  - increased use of color
- Animation
  - scripting
  - time sequences of data
  - animation of parameter changes such as isosurface value, cross-section position, or view point

Future Work - Long Term (cont.)

- Annotation
  - full control of size, color, placement, etc.
  - axis labels
  - color legends
  - time
  - titles
  - etc.
- Investigate Performance Enhancements
  For example:
  - geometry caching (storing geometries for each time step, so that they can be sequenced rapidly)
  - memory management to reduce paging (data compression might be used to reduce memory usage)
- Port Capabilities into AVS
- Incorporation of Imagery
- Improved Handling of Missing Data
- On-going investigation of extensions in areas such as new visual representations, virtual reality, and video
- Beta Release Through COSMIC & Unidata

PLANETARY DATA ANALYSIS AND DISPLAY SYSTEM: A VERSION OF PC-McIDAS

Dr. Sanjay S. Limaye
University of Wisconsin-Madison

August 12, 1992

PLANETARY DATA ANALYSIS AND DISPLAY SYSTEM

A VESION OF PC-MCIDAS

SANJAY S. LIMAYE
PRINCIPAL INVESTIGATOR

SPACE SCIENCE & ENGINEERING CENTER
UNIVERSITY OF WISCONSIN-MADISON
MADISON, WISCONSIN

LIMAYE@WISMACC
TEL: (608)262-9541
FAX: (608)262-5974

McIDAS

AN EVOLUTIONARY HARDWARE/SOFTWARE SYSTEM FOR EARTH ATMOSPHERIC DATA IN USE SINCE MID 1970'S.

AT PRESENT PRIMARILY SERVING THE METEOROLOGICAL COMMUNITY FOR RESEARCH, OPERATIONAL AND EDUCATIONAL APPLICATION AT NATIONAL (NOAA, NASA), INTERNATIONAL (CHINA, AUSTRALIA, SPAIN, ECMWF) AND EDUCATIONAL INSTITUTIONS (NCAR/UNIDATA)

PLANETARY VERSION AIMED FOR ANALYSIS OF PRIMARILY IMAGING DATA FROM SPACE MISSIONS

SUCH AS VOYAGER, PIONEER VENUS, MAGELLAN, HUBBLE, MARS OBSERVER, CASSINI

FLAVORS OF McIDAS

* McIDAS-MVS FOR MAINFRAMES
* McIDAS-OS2 FOR OS/2 OPERATING SYSTEM
* McIDAS-X FOR UNIX WORKSTATIONS (RISC-6000, SGI, SUN) WITH X-WINDOWS

PLANETARY DATA

EACH MISSION HAS HAD A UNIQUE INSTRUMENTS AND DATA FORMAT

IMAGING SENSORS OF WIDE VARIETY:

- SINGLE CHANNEL FRAMING CAMERAS (VIDICON-MARINERS, CCD-GALILEO)
- SINGLE OR MULTIBAND LINE SCANNERS (MARS OBSERVER)
- SPIN SCAN IMAGING (PIONEER VENUS, PIONEER JUPITER/SATURN)
- ARTICULATED MULTISPECTRAL SCANNING (NIMS)

DATA MEDIA:

- EDR TAPES (VOYAGER, MARINER, VIKING)
- CD-ROM'S (MAGELLAN, VOYAGER)
- FITS FORMAT (HUBBLE)
- TCP/IP (MARS OBSERVER)

PLANETARY PC-MCIDAS IMPLEMENTATION BEGUN UNDER MCIDAS-X

MCIDAS-X RELEASED BY SPACE SCIENCE & ENGINEERING CENTER IN APRIL 1992 (REQUIRED FOR RUNNING THE PLANETARY APPLICATIONS)

CODE COMPATIBLE WITH MCIDAS-OS2

PLANETARY CODE TO BE TESTED BY JPL CO-INVESTIGATORS BY END OF 1992

IBM RISC-6000 WITH X-STATIONS ACQUIRED FOR DEVELOPMENT

3-COLOR TASKS AWATING NEW HARDWARE THAT SUPPORTS X-WINDOWS

TASKS/STATUS:

SYSTEM LEVEL DEVELOPMENT

* EXPANDED DATA DESCRIPTOR BLOCK (DATA DIRECTORY STRUCTURE) - *DEVELOPED*
* MULTIBAND DATA DISPLAY AND PROCESSING - *UNDER TESTING*
* PROCESSING TRAIL - COMMAND INPUT (DEFAULT VALUES) INFORMATION - *PENDING*
* USER INTERFACE - *PENDING*

NEW SOFTWARE APPLICATIONS

* GENERALIZED NAVIGATION OF PLANETARY DATA - *PROGRESS*
* NEW ANALYSIS APPLICATIONS - *PROGRESS*
  - LOCATION SPECIFIC TIME SERIES ANALYSIS - *TESTING*
  - MULTISPECTRAL ANALYSIS - *DEVELOPED*
  - DATA VISUALIZATION - *PENDING*
  - EXPORT/IMPORT OF DATA IN STANDARD FORMATS - *PROGRESS*

DATA DESCRIPTOR BLOCK

WHAT IS IT?

- MCIDAS INTERNALLY HANDLES DATA IN TERMS OF "AREAS"
- EACH AREA HAS A DIRECTORY OR A "DATA DESCRIPTOR BLOCK"
- 64 WORDS USED FOR SYSTEM USE
- 128 WORDS FOR DATA NAVIGATION
- 512 WORDS FOR APPLICATIONS PROGRAM USE (DATA SPECIFIC) (MOONS, SPACECRAFT, SUN AND STAR POSITIONS)

WHY?

- ALLOWS EASY ACCESS TO KEY DATA FOR APPLICATIONS (USERS AND SYSTEM)
- KNOWN PLACE FOR DATA TYPE, CALIBRATION, NAVIGATION ETC.
- FOLLOWS THE DATA FROM MACHINE TO MACHINE AND TO ARCHIVE
- PROCESSING HISTORY WITH TIME AND DATE STAMP

MULTIBAND DATA DISPLAY AND PROCESSING

- PLANETARY AND EARTH DATA NOW OFTEN IS MULTISPECTRAL
- DATA STRUCTURES CAN HANDLE UP TO 32 BANDS
- MCIDAS IN THE PAST COULD DYNAMICALLY OVERLAY ONLY TWO CHANNELS
- HARDWARE AND SOFTWARE LIMITATIONS OF THE PAST NOW NO LONGER
- NEED THREE CHANNEL DYNAMIC COMPOSITE DISPLAYS WITH OVERLAY GRA PHICS CAPABILITY WITH INDEPENDENT FRAME (CHANNEL) CALIBRATION AND NAVIGATION

- PLANETARY COLOR IMAGES FROM THREE INDEPENDENTLY SHUTTERED MONOCHROME IMAGES WITH DIFFERENT NAVIGATIONS AND CALIBRATIONS
- NEED TO CREATE MULTIBANDED DATA STRUCTURE FOR 3-COLOR DISPLAY/ANALYSIS

Space Science & Engineering Center / University of Wicsonsin

Multispectral Analysis

- A generalized spectral classification (6 bands simultaneously) scheme developed to work under McIDAS-X.

Two versions:

- Iterative Self Organizing Data Analysis Technique (ISODATA)
  Initial arbitrary cluster means used to classify an
  image repetitively using minimum spectral distance
  to mean to assign classes to each pixel.
  Arbtitrary means shift to means of the clusters in data
  not spatially biased and better results than sequential
  clustering for data that are not normally distributed.
  Disadvantage that the range of input parameters that
  lead to convergence is narrow and simplistic criteria
  for cluster splitting and merging.

Limaye - 8

Space Science & Engineering Center / University of Wicsonsin

Transformed Divergence

Addresses the disadvantages of the first scheme

Improved splitting and merging of clusters

Results tested with ERDAS output --> as good or better analysis of planetary data

Limaye - 9

Space Science & Engineering Center / University of Wicsonsin

Optical Navigation

Voyager image pointing data contain roll angle errors

Use absolute reference for roll angle correction

- Read Hubble guide star files
- Detect stars (magnitude down to 10) in planetary images used for analysis
- Compute pointing correction

Technique successfully used on several hundred Voyager images to remove a + 0.85° error in roll angle!

Limaye - 10

EXPERIMENTER'S LABORATORY FOR VISUALIZED INTERACTIVE SCIENCE

Dr. Elaine Hansen
Colorado Space Grant Consortium

August 12, 1992

PROGRESS AND DIRECTION REPORT

EXPERIMENTER'S LABORATORY FOR VISUALIZED INTERACTIVE SCIENCE

Applied Information Systems Research Program

Workshop II

August 12, 1992

- **A group effort**

Colorado Space Grant Consortium:	*Elaine Hansen* *Allison Kipple* *Mike Porkert*
Laboratory for Atmospheric and Space Physics:	*Margi Klemp* *Eric Hillis* *Phil Evans*
National Center for Atmospheric Research:	*Joseph Klemp* *Bill Boyd* *Scott Davis*
Goddard Space Flight Center:	*Marti Szczur*
U C Santa Barbara:	*Jeff Star*

- **A slow start**
- **A group presentation**
  - Overview and User Interactions — *Elaine Hansen*
  - Usability Analysis and Software Design — *Margi Klemp*
  - The Rendering Package — *Bill Boyd*

PROGRAM GOALS

- Provide a capability that will help scientists of the '90's to interactively visualize data in order to better understand the *large, complex, diverse*, and *multidimensional* data sets of future challenges
- Provide an interactive visualization environment which supports science research *within and across NASA science disciplines*
- Provide a visualization laboratory that can be *easily used* and *easily tailored* by the scientists themselves to best fit their individual research problems and display preferences
- Enable general members of the Space Science Community to use advanced visualization tools *at an affordable price*
- *Capitalize* on existing information systems techniques, technologies, and tools

EXPERIMENTER'S LABORATORY FOR VISUALIZED INTERACTIVE SCIENCE

USABILITY ANALYSIS AND SOFTWARE DESIGN

Margi K. Klemp
Laboratory for Atmospheric and Space Physics
University of Colorado at Boulder

Applied Information Systems Research Program
Workshop II
August 12, 1992

PROJECT OBJECTIVES

- Develop a visualization and analysis application which will provide 3D surface and volumetric renderings with 2D graphics as required for quantitative analysis. It will run on a range of entry level to high end Unix workstations with no additional software requirements.
- Focus on usability
  - — User interviews throughout project
  - — Integrated design
  - — User testing with prototypes
  - — Iterative design
- Create general purpose tools to facilitate development of this project as well as other applications

USER SCENARIO

- Convert data to appropriate format for visualization application
- Process data (calibrate, filter, grid, average, etc.)
- Exploratory analysis:
  - — Visualize data qualitatively (image, 3D surface or volumetric rendering)

    User may be required to create polygons, set up color table, select light sources, etc. before rendering.
  - — Extract points, lines, cutting planes and volumes to get quantitative information
  - — Use interactive techniques to rotate, zoom, pan through data
  - — Relate data to other data sets by visualizing and interacting with several data sets simultaneously
  - — Manipulate data and visualize again
  - — Animate data to visualize changes with time
- Communicate and publish results
  - — Annotate data
  - — Create hard copy, videos, etc.

- **Supported by several members of Scientific Community**

Planetary	*Steve Lee*	PDS, LASP, CU
Solar	*Gary Heckman & Co.*	SEL, NOAA
	O. Richard White	HAO, NCAR, SOLSTICE, UARS
Earth Sciences	*Jeff Star*	UCSB
	Greg Scharfer	Snow & Ice Data Center/ EOS DAAC, NCAR
	Gary Rottman & Co.	SOLSTICE, UARS
Meteorology	*Joe Klemp & Co.*	NCAR

- **Progress**
  - User interviews
  - First round of requirements
  - Great interest in becoming test sites
- **Directions**
  - Additional Interviews
  - User evaluation / use of prototypes
- **Hopes**
  - ELVIS tools used by range of science users and disciplines

USABILITY CHALLENGES

- Integrate visualization, data processing, exploratory analysis, and communication needs into a single application
  - — Provide interface to external functions
  - — Design software to allow addition of new functionality (object-oriented approach)
- Provide a user interface which will hide the complexities of the hardware and software required to enable these functions (High level interface incorporating knowledge of user's domain)
- Allow the flexibility to make changes dynamically in the application and to add user specific functionality
  - — Programming languages are difficult to learn
  - — Spreadsheet paradigm has proven to be powerful and easy to learn
- Facilitate input of diverse data formats
- Provide annotation (automatic and direct manipulation)
- Develop direct manipulation interfaces for interaction with data

SOFTWARE DESIGN

- C++ object layer provides common programming interface to all software modules (PolyPaint, NetCDF, TAE+ and data objects)
- Spreadsheet engine used to propagate changes in cells attached to an object attribute (Based on Dr. Clayton Lewis' spreadsheet approach to interactive graphics)
- New graphics functionality will be created in object-oriented style (X widgets, TAE+ presentation types) to enhance reusability of code

PROJECT STATUS

- User interviews were conducted to gain understanding of the task
- Surveys of current visualization applications were made to determine their strengths and weaknesses
- An object-oriented design of the software was completed
- Objects were implemented in C++

  (significant enhancements were made to the color object to add flexibility and to hide the hardware dependencies on indexed or true color)
- PolyPaint was restructured to fit the object-oriented design
- The user interface was prototyped using TAE+

  (Some user testing and iterative design was completed)
- User interface code was written for initial "demonstration" version of application

FUTURE DIRECTIONS

- Complete Alpha version by end of year
- Continue user testing and design iteration
- Integrate spreadsheet engine
- Develop direct manipulation user interface as spreadsheet capabilities are available
- Add new graphics capabilities
  - — Color-filled contour plots
  - — Color-coded vector plots with curved vectors and user definable lengths
  - — Raster images
  - — XY graphs
  - — Histograms

FUTURE DIRECTIONS *(cont'd.)*

- Enhance TAE+
  - Add new presentation types based on new graphics capabilities
  - Incorporate spreadsheet engine in TAE+ Workbench to enable spreadsheet approach to programming user interface code
  - Use spreadsheet engine to create high level interface for window geometry management (resizing windows and objects in windows)
- Provide hooks for calling external functions
- Develop direct manipulation annotation editor
- Provide color hard copy interface
- Create interface to GL for hardware rendering

SAVS: A SPACE DATA ANALYSIS AND VISUALIZATION SYSTEM

Dr. Edward Szuszczewicz
Laboratory for Atmospheric and Space Science
Science Applications International Corp.

August 12, 1992

SAVS
A SPACE DATA ANALYSIS AND VISUALIZATION SYSTEM

SAIC
Laboratory for Atmospheric and Space Science

AVS, Inc.
Advanced Visual Systems

U of Md
Department of Physics and Astronomy

A Briefing Prepared By

E. Szuszczewicz, A. Mankofsky, P. Blanchard...SAIC
C. Goodrich, D. McNabb...U of Md
M. Moroh...CUNY D. Kamins...AVS, Inc.

August 11-13, 1992
AISRP Workshop, Boulder, Colorado

SAIC Laboratory for Atmospheric and Space Science McLean, VA 22102

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

INNOVATIVE VISUALIZATION SOFTWARE (AVS)

ADVANCED DATABASE TECHNIQUES

SET OF MATHEMATICAL, ANALYTICAL AND IMAGE PROCESSING TOOLS

STRONGLY DEVELOPED SENSE OF SCIENTIFIC REQUIREMENTS

SAIC Laboratory for Atmospheric and Space Science McLean, VA 22102

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

EXISTING NASA PLATFORMS	Interplanetary Medium, Near-Earth Magnetosphere
ISTP	The Sun, Interplanetary Medium, Outer and Inner Magnetosphere
SUNDIAL	The Global Ionosphere and Thermosphere
GRAND TOUR CLUSTER	Critical Magnetospheric Regions and Boundary Layers
TIMED	The Ionospheric-Thermospheric-Mesospheric System

SAIC Laboratory for Atmospheric and Space Science McLean, VA 22102

STATEMENT OF THE PROBLEM

Increased focus on large-scale system phenomena

The cross-disciplinary nature of many investigations

Higher data rates and projections of increased volumes of data

Enhanced measurement capabilities (need for cross-correlation of global images with "in situ" and ground-based observations)

Increasing number of large-scale 3-D numerical codes available as analytical tools for data synthesis and interpretation

SAIC Laboratory for Atmospheric and Space Science McLean, VA 22102

NEEDS OF THE PRACTICING SCIENTIST

Interactive data analysis and graphics *environment*

Ability to cross disciplinary boundaries with ease and understanding

Ability to "compress" data into a visually-organized form optimized for analysis and interpretation

Easy-to-use mathematical statistical and image processing tools

Tools to *obtain data sets from remote archives*

Access to empirical and numerical *model results* to correlate with the data and assist in data anlysis and interpretation

An *integrated* user-friendly *system they can afford*

SAIC Laboratory for Atmospheric and Space Science McLean, VA 22102

THE VISUALIZATION SYSTEM (AVS)

AVS IS DESIGNED FOR A DISTRIBUTED NETWORK ENVIRONMENT
...SINGLE SYSTEM OR A NETWORK OF SYSTEMS

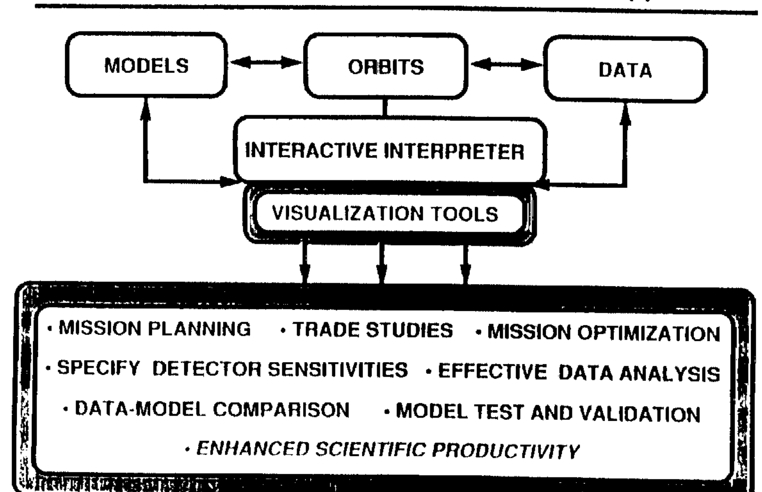
COMPLETE IMAGE DISPLAY CAPABILITIES:
...REAL-TIME PAN AND ZOOM
...ROTATION AND TRANSFORMATION
...FLIPBOOK ANIMATION
...SUPPORT FOR 8-BIT, 24-BIT AND FLOATING POINT IMAGES

IMAGE FILTERS INCLUDE:
...LOOK-UP TABLE OPERATIONS (PSEUDO-COLORING, HISTOGRAM BALANCING, DATA RESIZING, INTERPOLATION, CROPPING AND SAMPLING)

PROVIDES A VARIETY OF TOOLS FOR RENDERING VOLUME DATA; A REAL-TIME ISO-SURFACE GENERATOR; ETC.

SAIC Laboratory for Atmospheric and Space Science McLean, VA 22102

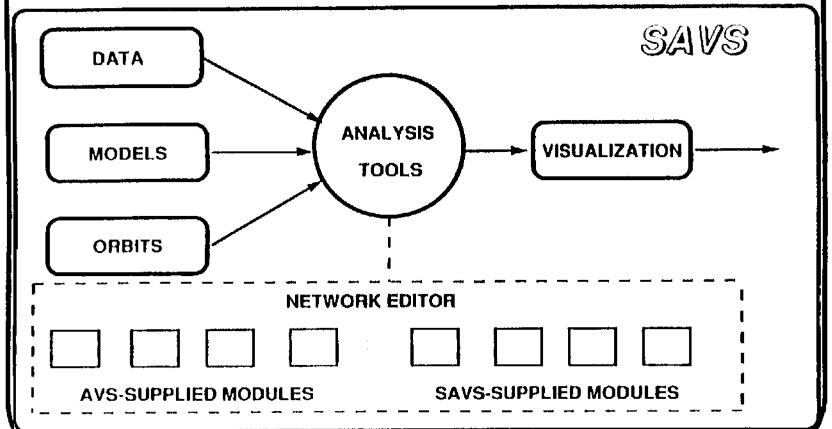
SAVS — *VISUALIZATION TOOLS An End-to-End Approach*



SAIC Laboratory for Atmospheric and Space Science McLean, VA 22102

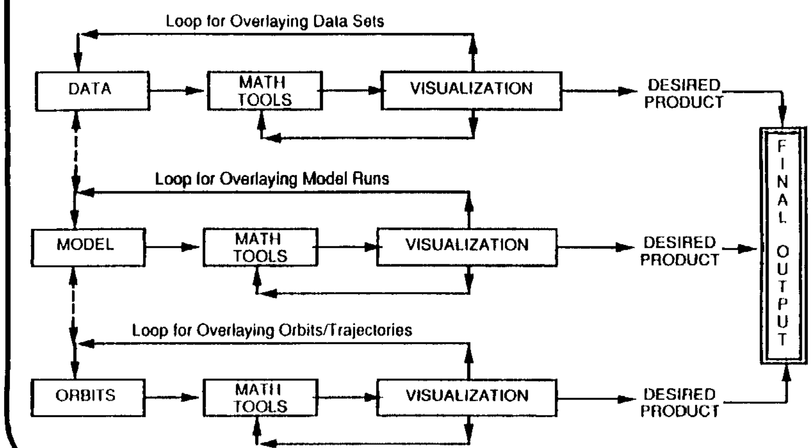
SAVS Design Goals — Ease & Functionality

- INTEGRATION OF VISUALIZATION WITH DATA MANAGEMENT AND ANALYSIS
- EXTENSIBILITY WITHOUT COMPLEXITY
- APPLICATION DEVELOPMENT WITH THE SCIENTIST AS THE END USER

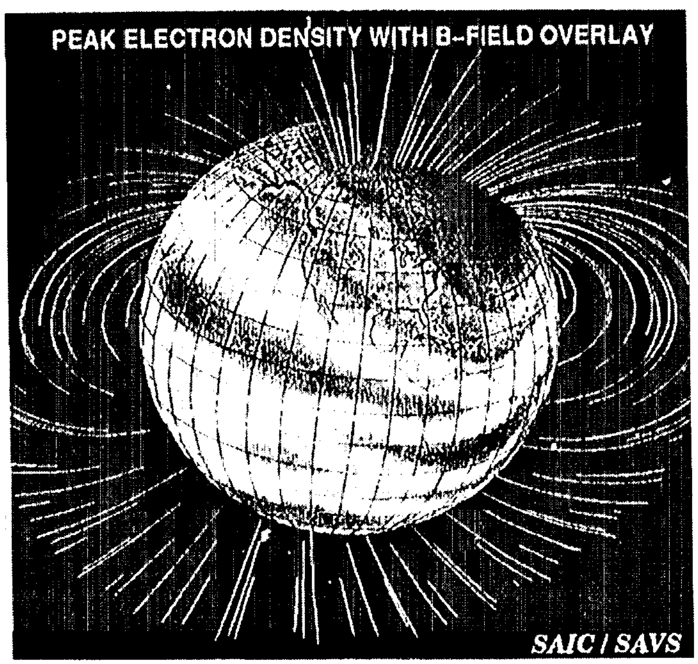


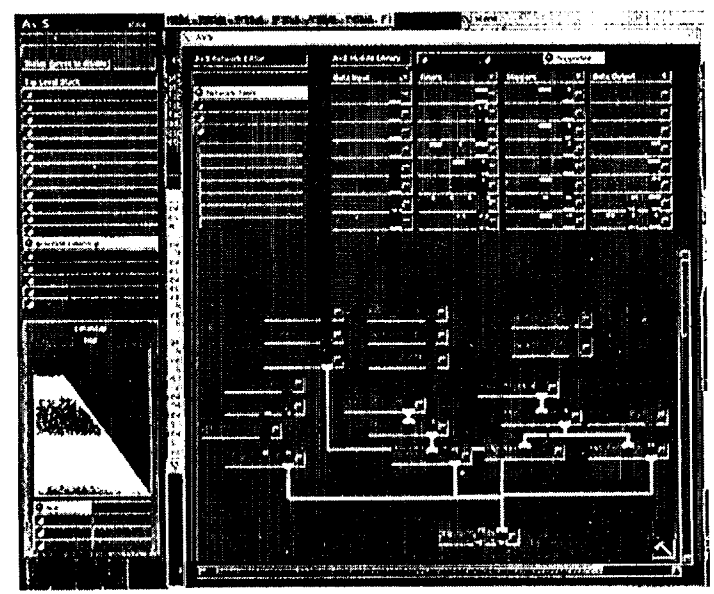
SAIC Laboratory for Atmospheric and Space Science, McLean, VA 22102

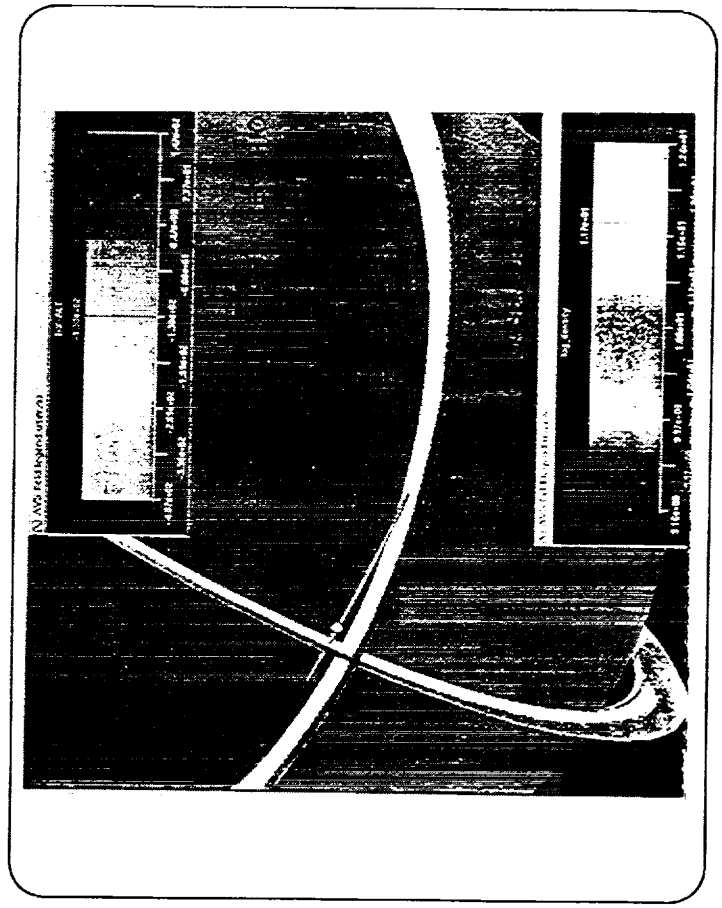
SAVS: Data Analysis & Visualization — Interactive Functionality

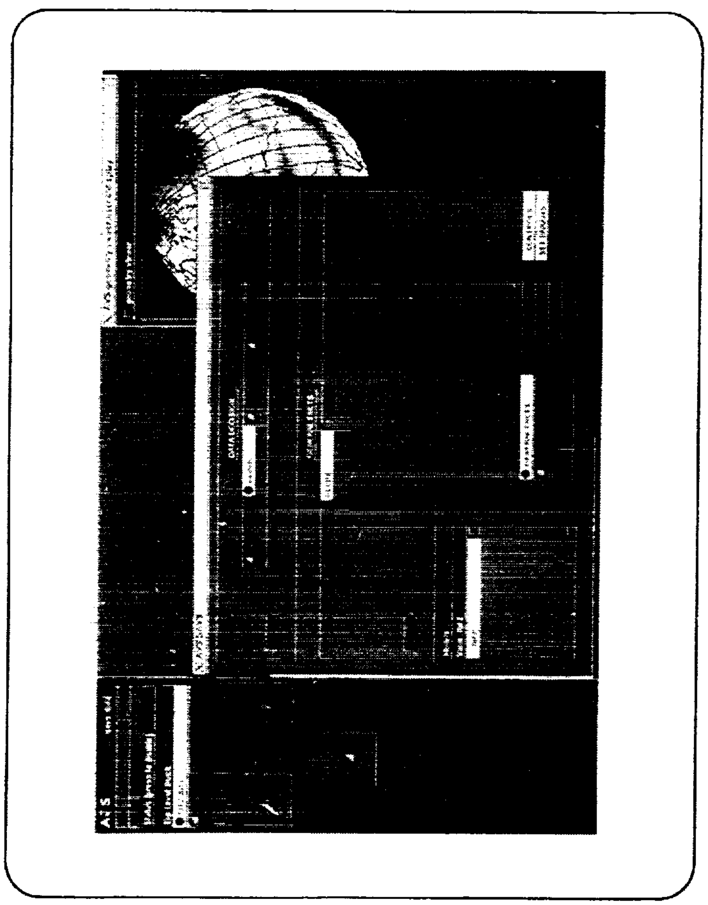


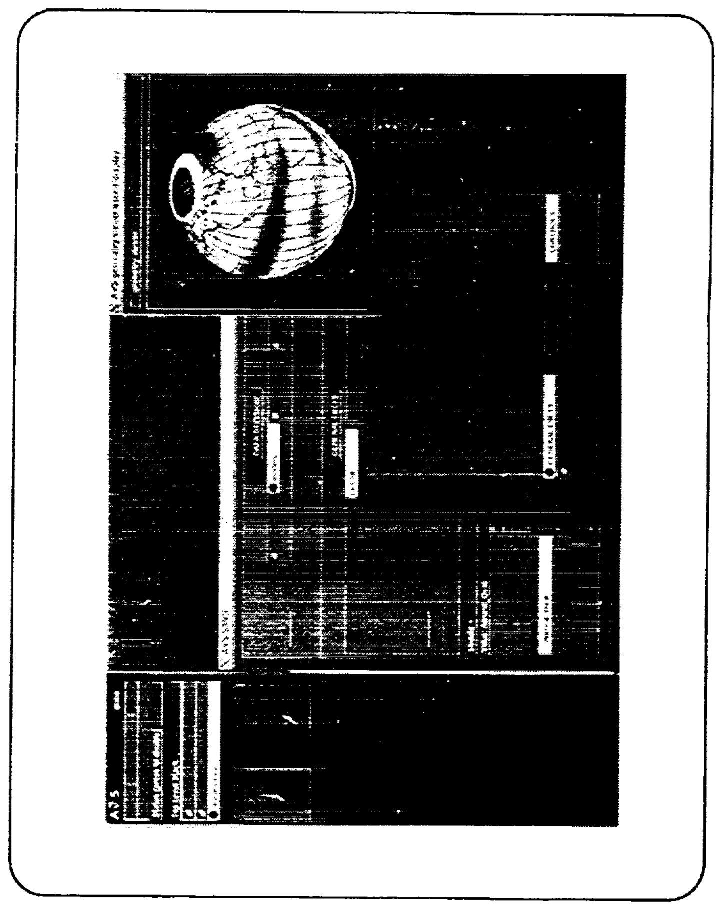
SAIC Laboratory for Atmospheric and Space Science, McLean, VA 22102

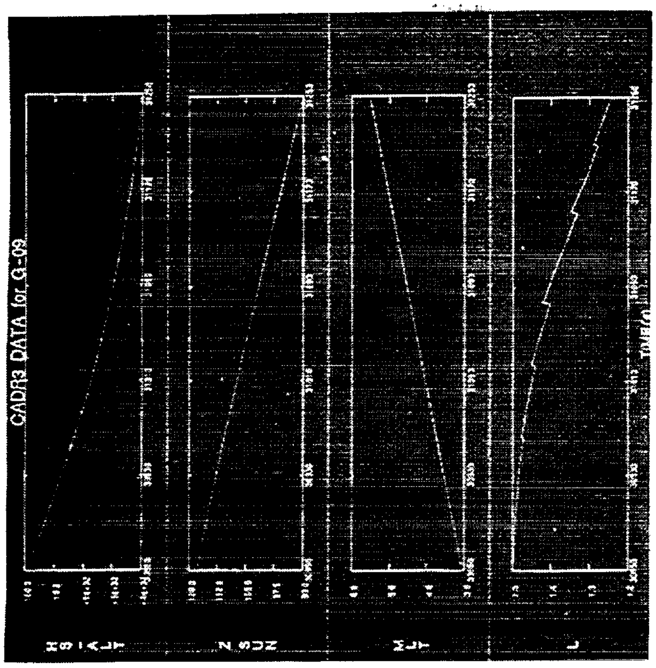


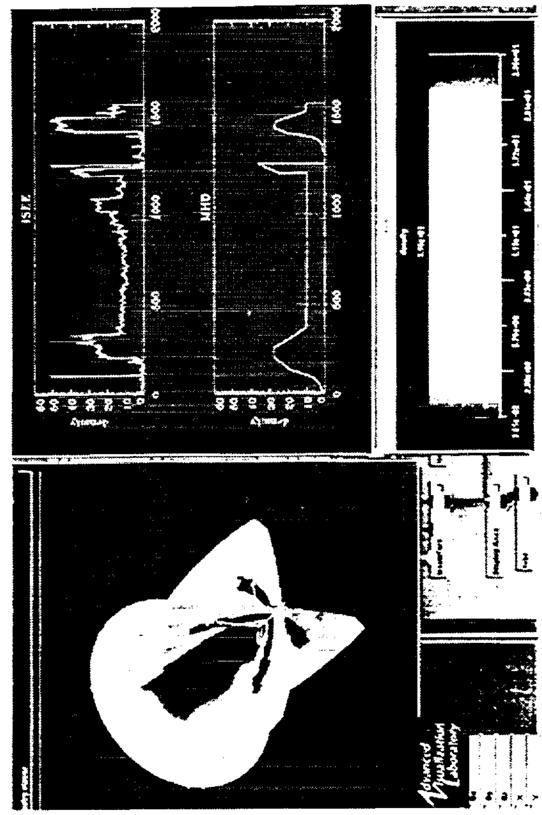


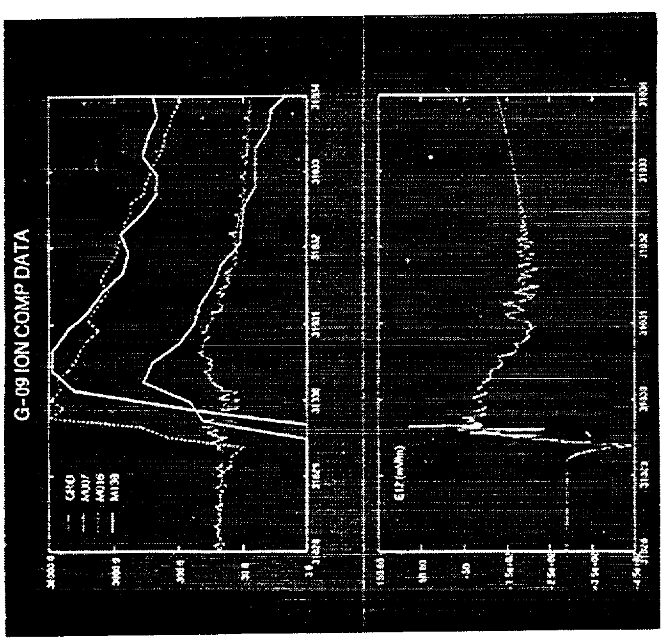


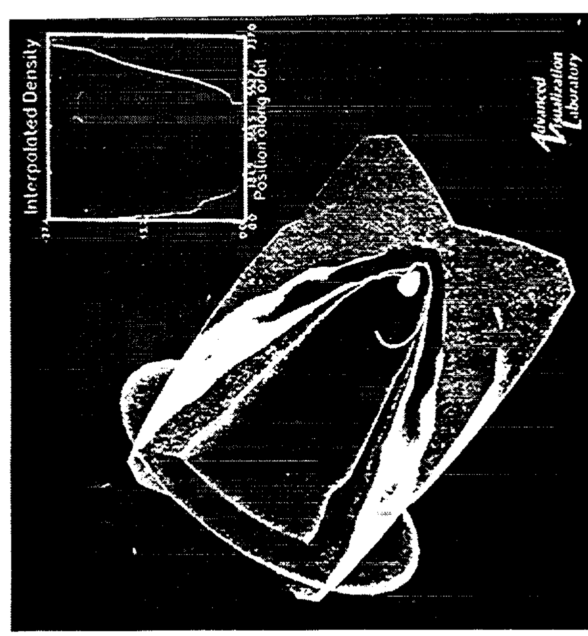





GADR3 DATA for G-09


ISEE
MHD
Advanced Visualization Laboratory


G-09 ION COMP DATA


Interpolated Density
Position along orbit
Advanced Visualization Laboratory

SAVS: Data Analysis & Visualization — First Year Accomplishments

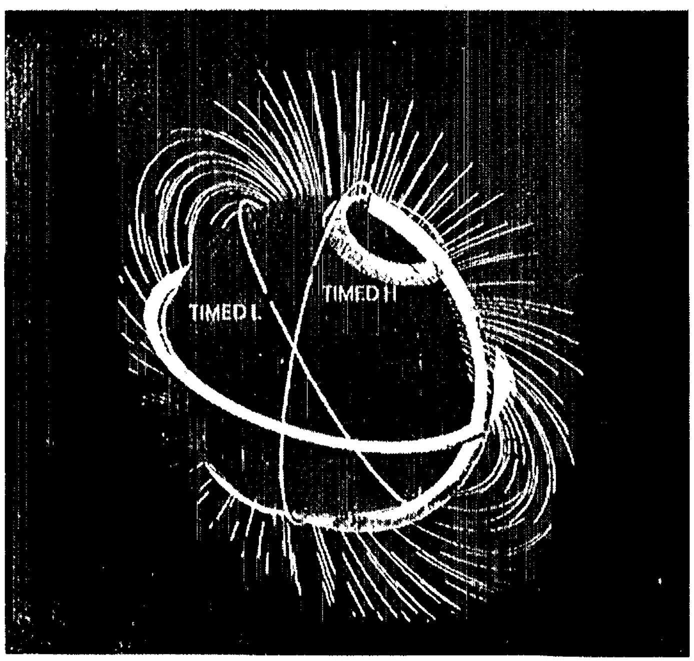
- AVS ported to lower-end platforms
- Customized AVS interface to NASA applications of 1-, 2-, and 3-d displays
- Developed extensible user-friendly architecture
- Developed data and model interface modules
- Implemented basic mathematical and statistical functions
- Began the development of hooks for an interactive interpreter
- Implemented general calls and execution modules for a number of large-scale models
- Tested system on CRRES and ISEE orbits and local data bases
- Initiated plans for remote data access capabilities

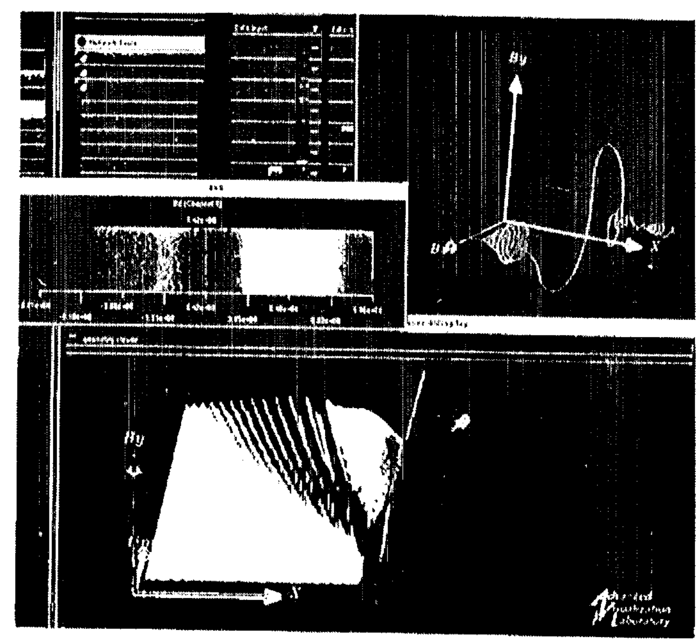
SAIC Laboratory for Atmospheric and Space Science McLean, VA 22102

SAVS: Data Analysis & Visualization — Second Year Plans

- Continue updating and upgrading screens, interfaces and overall architecture for ease, simplicity, and extensibility of the SAVS system
- Continue to develop general purpose "hooks" for relevant public domain codes (e.g., MSIS, TIEGCM, etc.)
- Develop recipes for new user applications
- Develop and integrate generic Remote Procedure Call capabilities
- Develop remote data access capability (i.e., develop set of directories, and staging, browsing and extraction tools)
- Test system capabilities for remote access and handling on NASA data subsets (candidates include: DE RIMS, PWI, and EICS; IMP-8 IMF, and ISEE)
- Develop and test generalized SAVS data input modules
- Develop and test generalized interpolation modules (e.g., arbitrary 3-d onto 1-d)
- Develop and test integrated SAVS/PV-Wave link to include binary direct memory transfer and bi-directional data transfer

SAIC Laboratory for Atmospheric and Space Science McLean, VA 22102





NASA SUPPORTED ADVANCED VISUALIZATION TECHNIQUES

Dr. Bill Hibbard
University of Wisconsin-Madison

August 12, 1992

NASA supported advanced visualization techniques

University of Wisconsin-Madison
Space Science and Engineering Center

Bill Hibbard and Brian Paul

UNIX McIDAS
the initial development of McIDAS-X

VIS-5D
VISualization of 5-Dimensional data sets

VIS-AD
VISualization for Algorithm Development

VIS-5D

VIS-5D is used to visualize large output data sets (10 ^ 8 grid points) from numerical weather models.

Greg Tripoli and his students have demonstrated that using VIS-5D for routine diagnostics of their UW-RAMS model has a real impact on the quality of science.

Used by scientists at UW, at NASA/MSFC, at NASA/GSFC, and at numerous other sites.

Runs on SGI, IBM RISC and Stardent.

Available as freeware by anonymous ftp.

VIS-AD

A visual laboratory for experimenting with algorithms for extracting useful information from remote sensing data.

Too much EOS data for people to be in the processing loop: one terabyte per day =

one million 1K x 1K 8-bit images per day.

EOS data processing must be automated.

VIS-AD puts visualization where the people work - developing algorithms for automatic processing.

VIS-AD is like an interactive debugger that produces visualizations of data objects instead of just printing their values - to help us see high-level behavior rather than just low-level bugs.

The user can:

- develop and modify an algorithm whose data objects have user-defined data types
- control execution by single steps or setting breakpoints
- view any combination of algorithm data objects in common user-defined frames of reference

All highly interactively in an integrated environment.

VIS-AD can also be viewed as a McIDAS macro language where the user can:

- invent data structures as needed
- display any data object simply by pointing and clicking at its name
- invent display frames of reference for:
  spatial display
  time series
  thermodynamic diagram
  scatter diagrams
  etc.

VIS-AD can access McIDAS data structures.

LINKWINDS: THE LINKED WINDOWS INTERACTIVE DATA SYSTEM

Dr. Allan S. Jacobson
Jet Propulsion Laboratory

August 12, 1992

JPL LinkWinds

The Linked Windows Interactive Data System

Principal Investigator

Allan (Bud) Jacobson - Jet Propulsion Laboratory

Investigator Team	*Developer Team*
Jet Propulsion Laboratory: Mark Allen, Ronald Blom, Donald Collins, Lee Elson	Jet Propulsion Laboratory: Andy Berkin, Bonnie Boyd, Martin Orton, Mitch Wade
San Diego Supercomputer Center: Phil Mercurio	San Diego Supercomputer Center: Jim McLeod

JPL LinkWinds

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.

JPL LinkWinds

The Linked Windows Interactive Data System

System Description

1. 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.

2. A standard graphical user interface with additional data-linking rules. Results in an interface which is highly interactive, intuitive and uniform across all applications.

3. Implemented on an object-oriented programming model, with "links" establishing message flow path. The underlying command language (Lynx) based upon scheme.

4. A multi-user science environment (MUSE) requiring minimum band-width and useful for cooperative scientific research, remote tutorials and development feedback.

JPL LinkWinds

The Linked Windows Interactive Data System

Development Approach

1. Don't develop solutions in search of problems.

2. Work with users and developers in a tight loop throughout the development process.

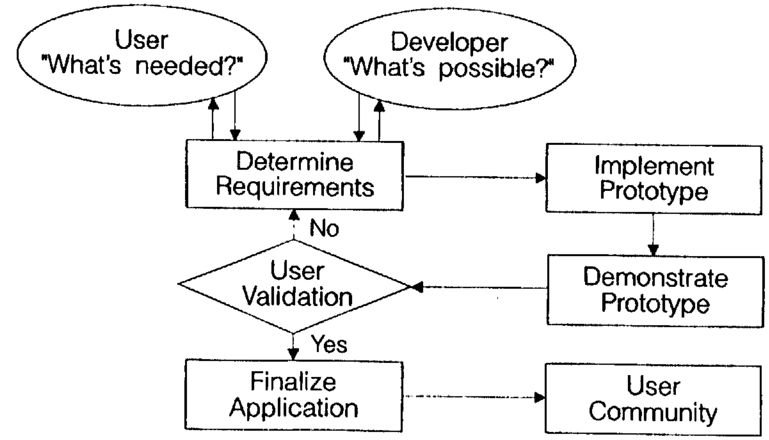
3. Employ an incremental development process using rapid prototyping of applications.

4. Provide software and workstations to collaborating users to stimulate product use in research activities.

JPL LinkWinds

The Linked Windows Interactive Data System

Application Development Cycle



JPL LinkWinds

The Linked Windows Interactive Data System

User Interface Design Philosophy

1. Users are impatient and want to get started quickly on productive work. They are discouraged by large manuals.

2. Users learn from self-initiated exploration, making mistakes and correcting them.

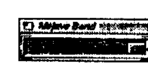
3. Users refer to documentation only when the software doesn't conform to their expectations. Then they skip around in manuals or on-line help to find the answer to current problem.

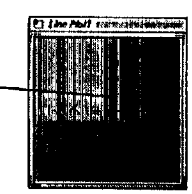
(Ref: Marc Rettig, CACM, Vol. 34, 19, July 1991.)

JPL LinkWinds

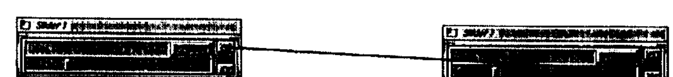
The Linked Windows Interactive Data System

Linking Rules

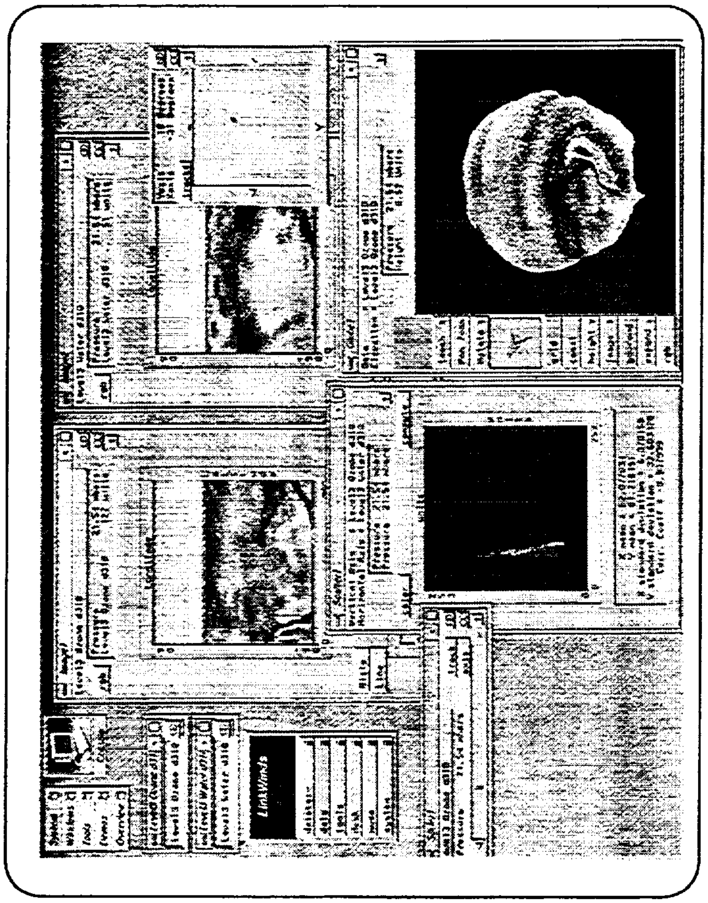


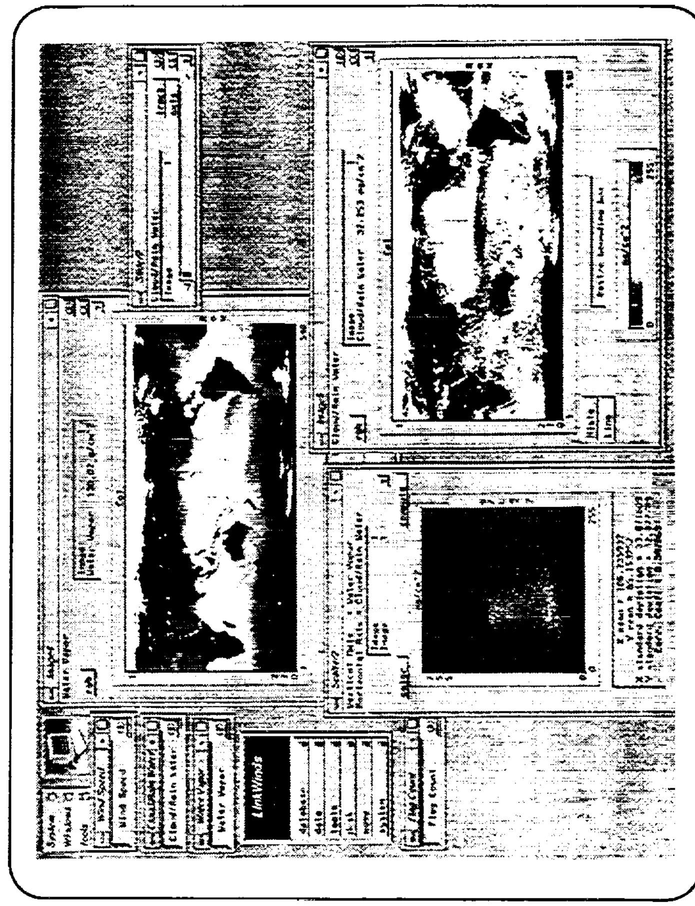


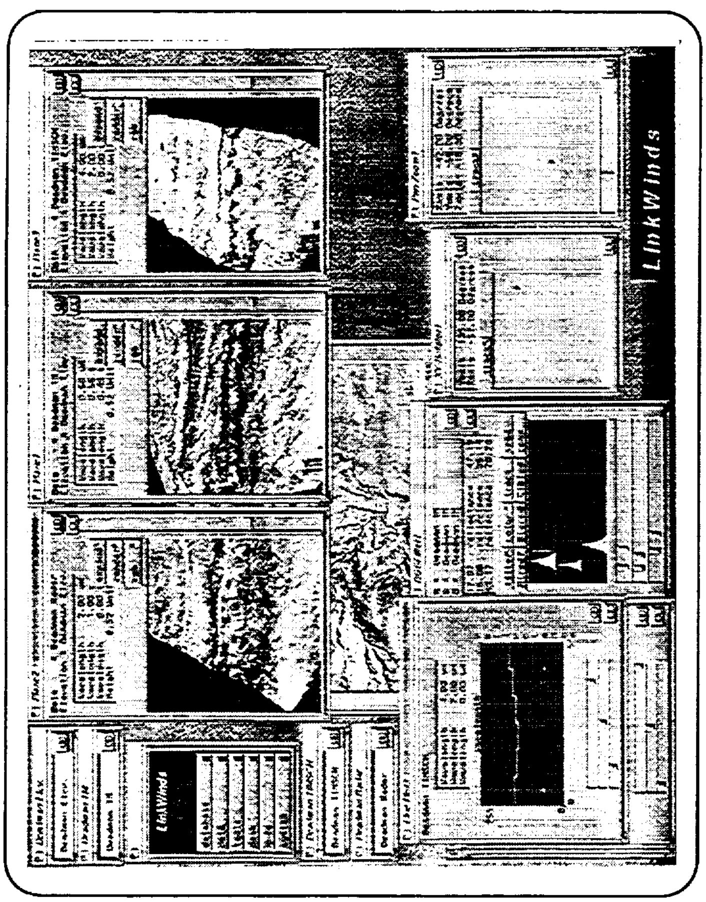
1. If an empty window appears, link a data object to it.

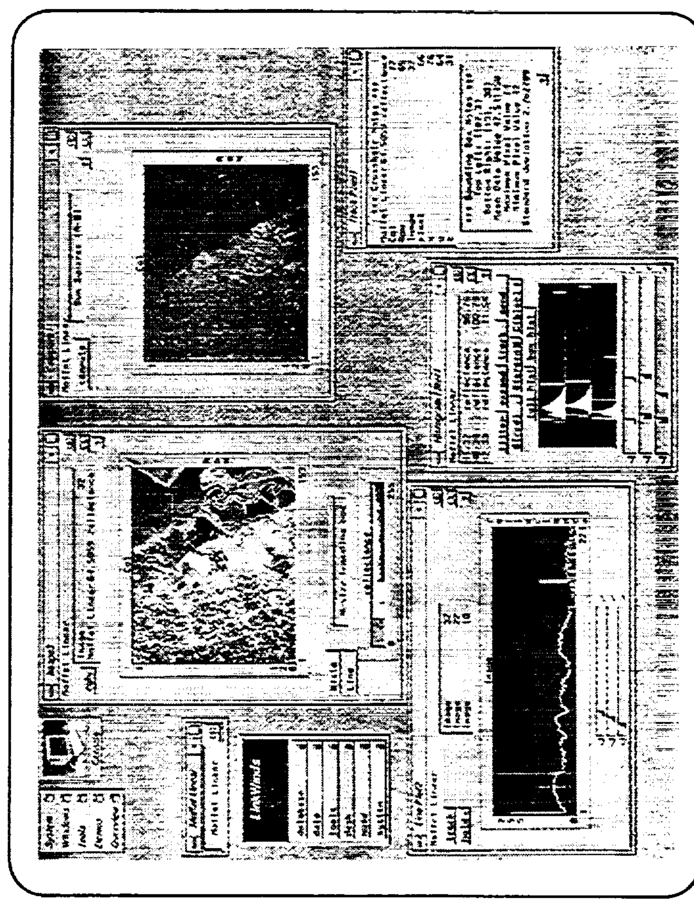


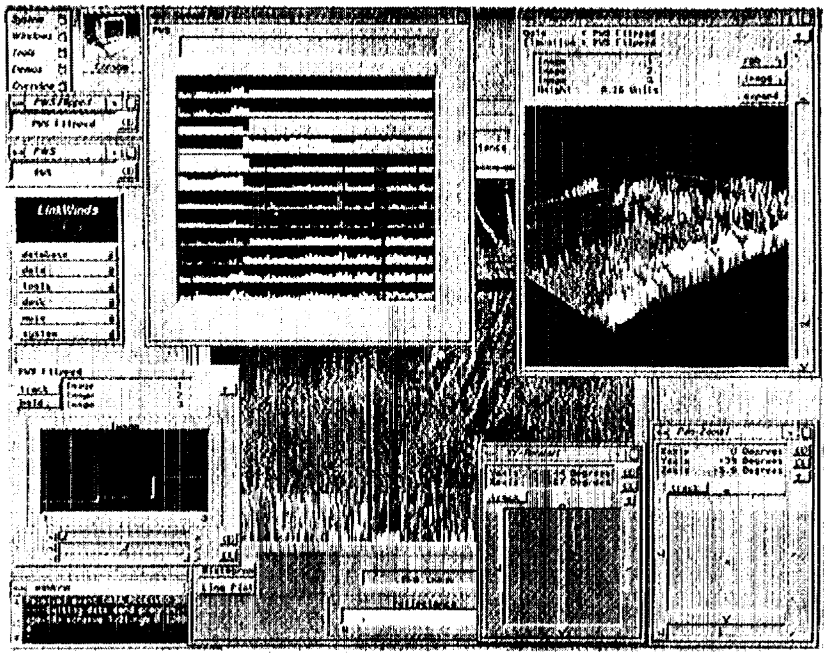
2. Link control symbols to other windows in order to manipulate them.






LinkWinds





JPL **LinkWinds**

The Linked Windows Interactive Data System

Sample Files for Ingesting Databases

The following three files are edited to load new databases into LinkWinds:

1. lw.config - The LinkWinds initializing configuration file. At the top are added the paths to any directories which contain the users data files.

2. Databases - A file which resides in one of the paths in lw.config. It lists all of the .db files which contain data file description and metadata information.

3. Menuname.db files - A file containing descriptions of the database files and any appropriate metadata.

JPL **LinkWinds**

The Linked Windows Interactive Data System

Database Interface

1. Currently standardized on Hierarchical Data Format (HDF) created at University of Illinois, Champagne/Urbana, NCSA.

2. Other data formats are accessible via interaction with DataHub.

3. User integrates database files into system by editing three text files containing data description and metadata. Templates are provided.

JPL **LinkWinds**

The Linked Windows Interactive Data System

Future Plans

1. Port LinkWinds to other unix platforms.

2. Expand standard input data formats via interaction with DataHub, and develop applications for visual data selection and subsetting. Simplify, simplify simplify.

3. Affect a wider distribution of LinkWinds.

4. Expand MUSE capabilities.
   Session Management, Floor Management, Telepointers, Network clipboard

5. Implement an applications generator to support user application development.

LinkWinds: An Approach to Visual Data Analysis

Allan S. Jacobson
Jet Propulsion Laboratory
California Institute of Technology
4800 Oak Grove Drive
Pasadena, CA 91109

April 4, 1992

INTRODUCTION

Modern space, sensor and computer technologies have made it possible to understand the Earth and its environs as never before. Population pressures and modern technology have also tended to make it imperative that we do so. To accomplish this, tremendous masses of data must be gathered, transported, stored and understood. To address some of these topics, a program of research is being conducted into the application of computer graphics to the problems of quickly exploring and analyzing very large amounts of scientific or engineering data. The objectives of the program are (1) to develop a software environment which will support the rapid prototyping of visual data analysis applications, while at the same time maintaining the high level of performance necessary for interactively manipulating graphical displays; (2) to develop a user interface that is truly intuitive and easy to learn and allows quick access to the software for the novice as well as the advanced user; (3) to provide a suite of sample applications which are useful across a variety of scientific disciplines; and (4) to provide tools to support user development of applications for this environment.

LINKWINDS

The Linked Windows Interactive Data System, or LinkWinds, version 1.3 is a prototype product of this research effort. In compliance with our research objectives, it is a visual data analysis/exploration system designed to rapidly and interactively investigate large multivariate and/or multidisciplinary data sets to detect trends, correlations and anomalies. LinkWinds is an integrated multi-application execution environment with a full graphical user interface (GUI). The system, operating under Unix, is based on an object-oriented programming model and is implemented in the C-language. For its graphical user interface and graphics support software, it draws upon the Silicon Graphics Inc (SGI) GL-library, and presently runs only on workstations supporting this library. This includes all SGI workstations, and those of other manufacturers who have licensed and support the GL-library.

Individual tools and data sets are coded as objects, each occupying a window on the LinkWinds screen, and communicating with other objects through a message passing protocol. The objects or windows, containing data displays and controls for manipulating these displays, can be linked or unlinked at the discretion of the user. The act of linking the windows together sets up one-way message paths. This data-linking paradigm makes the system perform much like a graphics spreadsheet, and as in a spreadsheet, is a powerful way of organizing the data for analysis while at the same time providing a natural and intuitive interface. Data-linking, and its user interface implications are discussed below.

Messages generated by LinkWinds objects are recorded as program statements in an underlying language called Lynx, which is based upon Scheme, a dialect of Lisp. A modified Scheme interpreter is included in the system. The message passing characteristics are the basis for two key LinkWinds functions. The first of these is the maintenance of an internal journal of all user originated commands executed by the environment. This file can be saved at anytime through a menu option. The record can then be replayed at the initiation of subsequent LinkWinds sessions, allowing the user to draw upon a previous layout of LinkWinds applications and links, or repeat a full analysis session.

The second function based upon the Lynx message passing protocol is the multi-user science environment (MUSE) which provides a method for multiple LinkWinds systems to communicate via networks. Using menu options, users remotely separated can connect to one another, and by also establishing a telephone voice connection, can cooperatively view and manipulate their data. A successful connection requires that each user be executing LinkWinds and that each has access to the data sets that are being analyzed. This is normally arranged by transporting the data sets to each user prior to the collaborative sessions. The MUSE capability is also used to give tutorials over the network to new users unfamiliar with LinkWinds, and to allow users to demonstrate recommendations for application changes or to point out bugs.

Hard copy of the LinkWinds displays are provided by function keys on the keyboard. Placing the cursor in a window, and pressing F1 produces an image of a window's contents; pressing F2 saves the complete window; and F3 saves the full screen. The figures shown were obtained in this manner.

DATA-LINKING AND THE USER INTERFACE

In addition to the normal GUI functions provided by the windowing environment, dynamic manipulation of graphs and images is facilitated through the data-linking paradigm. Data-linking can be understood in the context of a spreadsheet, where cells containing numbers are linked to other cells. Formulae are associated with each cell, so that when a number changes, all cells linked to the changed cell recalculate their values and update. LinkWinds does the same thing, but in a graphics environment where the rigid grid structure gives way to free form, and a cell can translate, for instance, into a slider or large scale number arrays such as images.

The user interface based upon the data-linking paradigm is one of the most distinguishing features of LinkWinds. It evolved from a desire to create a truly easy to learn and intuitive user interface. We are guided by the principle that users are impatient and want to get started on productive work as quickly as possible. Therefore, an interface was needed which can be learned by exploration, and which conforms to user expectations as they work with it.

Data-linking is affected through two icons. The link icon is a button displaying two interlocking rings, while the unlink icon displays two rings that are separated. Each object on the screen has either a single link button, or the full set of link and unlink buttons. The presence of a single link button indicates a data object, while the presence of the pair indicates applications with control functions. To perform a link, the cursor is placed on the appropriate button, and a "rubber band" is dragged out and dropped into the application to be linked. To break the link, the same thing is done using the unlink button. There are two simple rules to follow in applying the linking paradigm.

1. When as a result of menu selections an empty window appears on the screen, put data in it. This is done by linking a data object into the window.

2. When an object with the pair of link symbols appears, exercise its control function by linking it with any application object.

DATABASE INTERFACE

The current version of LinkWinds accepts data in the 8-bit raster Hierarchical Data Format (HDF) created and supported by the National Center for Supercomputing Applications (NCSA) at the University of Illinois, Champagne/Urbana. A 2D data file is a single image, while a 3D data file is a sequence of images. The data ingestion is controlled by two text files which are generated by the user. Sample versions of each of these files are provided and serve as self-explanatory templates. The first is a file which lists all of the databases of interest to the user. Those listed appear in the top level "Databases" menu. The second file is a data description containing the filenames of all data sets to be associated during the analysis, the number of axes and their names, any metadata needed to translate axis values to numbers meaningful to the data, etc. Data sets listed in these files appear in the top-level "Data" menu. A palette may be included in the HDF file if the user desires, or it may be defined in the data description file. Colors can also be assigned during a data analysis session using a color management tool provided. A much greater range of data formats will be acceptable to LinkWinds in future versions.

APPLICATIONS OVERVIEW

A suite of applications useful across many disciplines has been developed for the LinkWinds environment. Figure 1 shows a typical session to explore a data set collected by the Microwave Limb Sounder (MLS) currently in orbit aboard the Upper Atmospheric Research Satellite (UARS). The LinkWinds top-level menu is shown on the left and data objects, with their single link buttons, are in the upper left-hand corner. In this case, the data displayed are ozone and water vapor. The window entitled Image1 contains a slice of the data at an altitude of 21.54 mbars, as selected by Slider1 which is linked to it. Image2 is also linked to Slider1 and shows the water vapor at the same altitude. The southern hemisphere ozone hole is shown at the lower left of Image1, and a high value of the water vapor is shown in a corresponding location in Image2. The anticorrelation is shown in Scatter1, where the points shown come from the bounding box shown in Image1, controlled by linking Image1 to Scatter1. The ozone data are displayed in Globe1 as a height field rendered on a sphere. The ozone hole associated with the south pole is clearly seen. Slider1 also controls the depth of this display, and the height scale is controlled by a vertical slider along the right side of the window. Pan/Zoom and 2-Axis Rotator controls are also linked to Globe1 through the Animator. The Animator makes it possible to select a starting set of control settings, and an ending set. Then it will automatically record and save the number of frames selected with its slider, resulting in an animation of the sequence of settings from start to end. This sequence can be replayed later.

Figure 2 shows coregistered data collected in the Deadman Butte area of Wyoming. Line Plot# applications show spectral profiles for each of the data sets, and is controlled by a cross-hair on Image# dynamically associating a spectrum with each point on the terrain. Sliders on Line Plot# allow the selection of three of the channels to be colored red, green or blue, and RGBFilter# permits interactive color stretching of each of these channels. These also control the colors used for the perspective renderings of the data in Plane#. Thematic Mapper (TM), Quad-pole Synthetic Aperture Radar (SAR) and Thermal Infrared Multispectral Scanner (TIMS) data are superposed on an elevation map of the terrain in Plane1, Plane2 and Plane3, respectively. These in turn are oriented in unison by the Pan/Zoom and 2-Axis Rotator controls.

ACKNOWLEDGMENTS

The current LinkWinds development team members are Andy Berkin, Bonnie Boyd, Martin Orton and Mitch Wade. Alumni who made important contributions are Brian Beckman and Leo Blume. LinkWinds couldn't have reached its current state of usefulness without the close collaboration of users who are applying the software in their research and providing comments and suggestions vital to its development. These scientists are Mark Allen, Kathi Beratan, Ronald Blom, Donald Collins, Robert Crippen, Lee Elson and Andy Tran. We extend our thanks to Joe Bredekamp and his staff in NASA Code SMI, whose sponsorship of this work has made it possible.

For additional information, contact the author at 818-354-0693, or any of the team members at 818-354-0728.

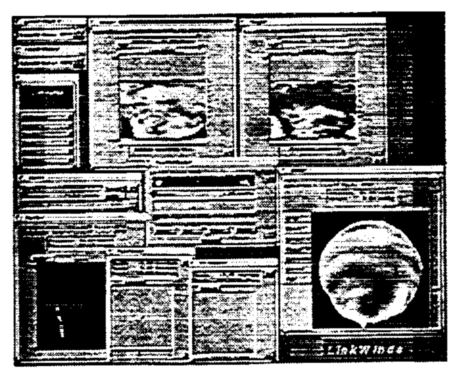


Figure 1 - LinkWinds session to explore upper atmospheric ozone and water vapor measured by the Microwave Limb Sounder aboard UARS.

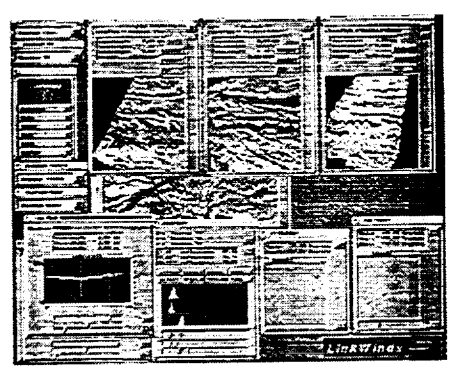


Figure 2 - Coregistered spectral data gathered in the region of Deadman's Butte in Wyoming. Several spectral bands are shown rendered in perspective.

DATAHUB: KNOWLEDGE-BASED SCIENCE DATA MANAGEMENT

Dr. Tom Handley
Jet Propulsion Laboratory

August 12, 1992

ASPRS/ACSM/RT 92 CONVENTION

DataHub: Knowledge-Based Science Data Management

JPL

Thomas H. Handley, Jr.
Y. Phillip Li
Allan S. Jacobson
An V. Tran

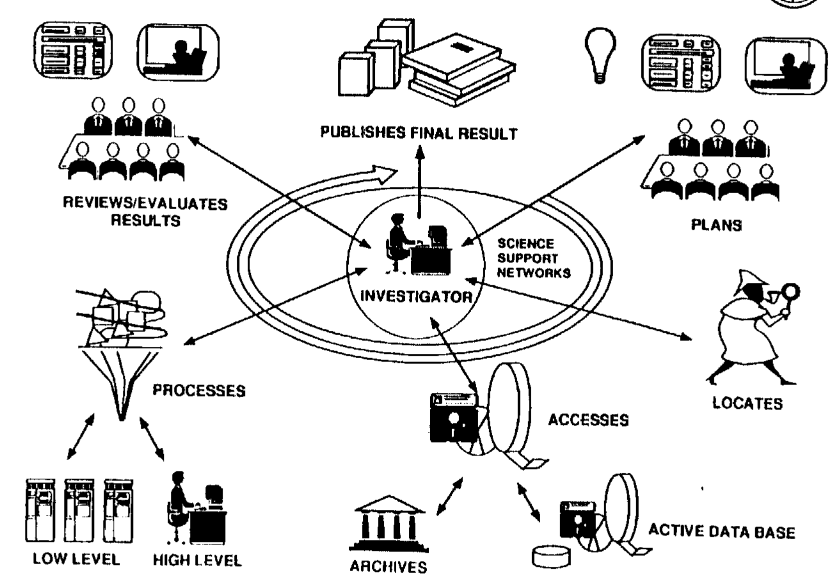
August 3-8, 1992

Jet Propulsion Laboratory
California Institute of Technology
4800 Oak Grove Drive
Pasadena, California 91109

JPL

ASPRS/ACSM/RT 92 CONVENTION

SCIENCE DISCOVERY CYCLE



JPL

ASPRS/ACSM/RT 92 CONVENTION

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

ASPRS/ACSM/RT 92 CONVENTION

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

ASPRS/ACSM/RT 92 CONVENTION

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

ASPRS/ACSM/RT 92 CONVENTION

JPL

DataHub

- DataHub – A VALUE-ADDED, KNOWLEDGE-BASED SERVER BETWEEN THE DATA SUPPLIERS AND THE DATA CONSUMERS
- SCIENTIFIC DATA MODELS
    - DATA DRIVEN ANALYSIS
    - DATA TRANSFORMATIONS
    - DATA SEMANTICS
    - ANALYSIS-RELATED KNOWLEDGE ABOUT DATA
    - DATA DISCOVERY, INGESTION, EXTRACTION, ...
    - SELF-DESCRIBING DATA STRUCTURES
- INTELLIGENT ASSISTANT SYSTEM(S) WITH SOME KNOWLEDGE OF DATA MANAGEMENT AND ANALYSIS BUILT-IN
- USE OF MATURE EXPERT SYSTEM TECHNOLOGY TO AID EXPLORATORY DATA ANALYSIS, I.E. EXPERT SYSTEMS, NEURAL NETS, CLASSIFICATION SYSTEMS
- CAPTURE AND ENCODE KNOWLEDGE ABOUT THE DATA AND THEIR ASSOCIATED PROCESSES. ENCODE SCIENTIFIC KNOWLEDGE INTO THE ROUTINES, PROCESSES AND PROCEDURES
- PROVIDE DATA MANAGEMENT SERVICES TO EXPLORATORY DATA ANALYSIS APPLICATIONS(S) I.E. LinkWinds

ASPRS/ACSM/RT 92 CONVENTION

JPL

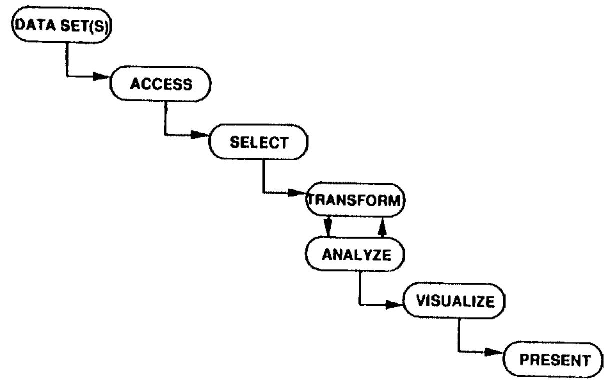
DataHub

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

ASPRS/ACSM/RT 92 CONVENTION

JPL

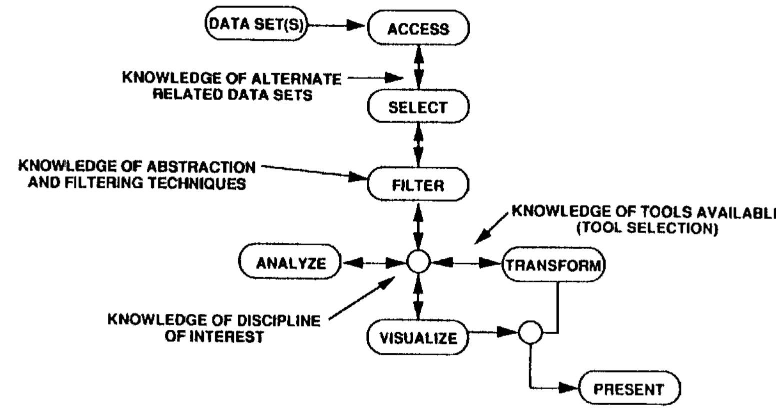
TRADITIONAL DATA ANALYSIS



ASPRS/ACSM/RT 92 CONVENTION

JPL

KNOWLEDGE-BASED DATA ANALYSIS AND VISUALIZATION



ASPRS/ACSM/RT 92 CONVENTION

JPL

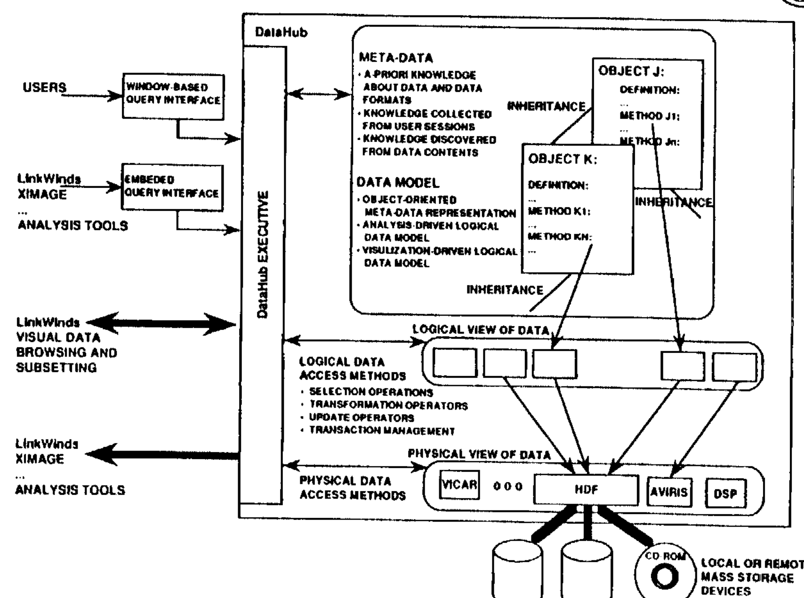
LinkWinds THE LINKED WINDOWS INTERACTIVE DATA SYSTEM

- 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

ASPRS/ACSM/RT 92 CONVENTION

JPL

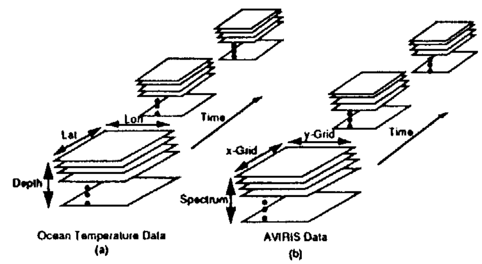
FUNCTIONAL ARCHITECTURE



ASPRS/ACSM/RT 92 CONVENTION

JPL

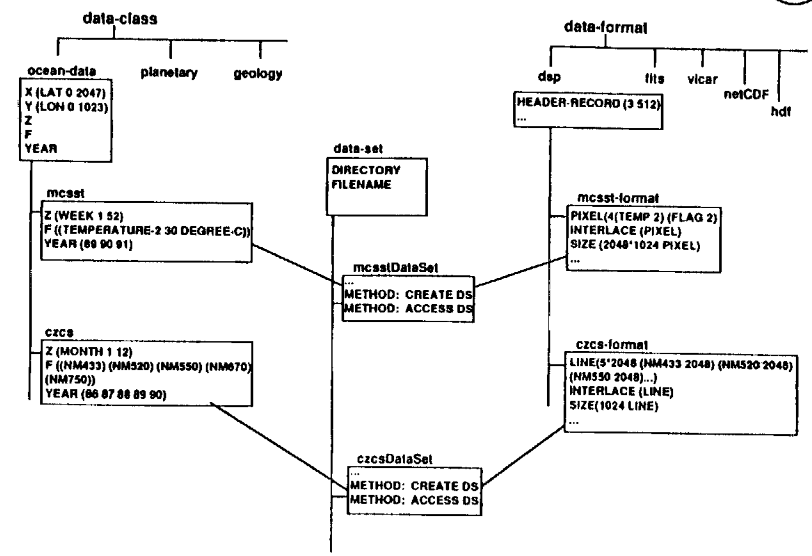
UNIFORMLY GRIDDED DATA



ASPRS/ACSM/RT 92 CONVENTION

JPL

DATA MODEL

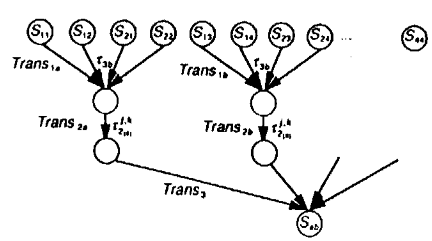


ASPRS/ACSM/RT 92 CONVENTION

JPL

TRANSACTION SPECIFICATION

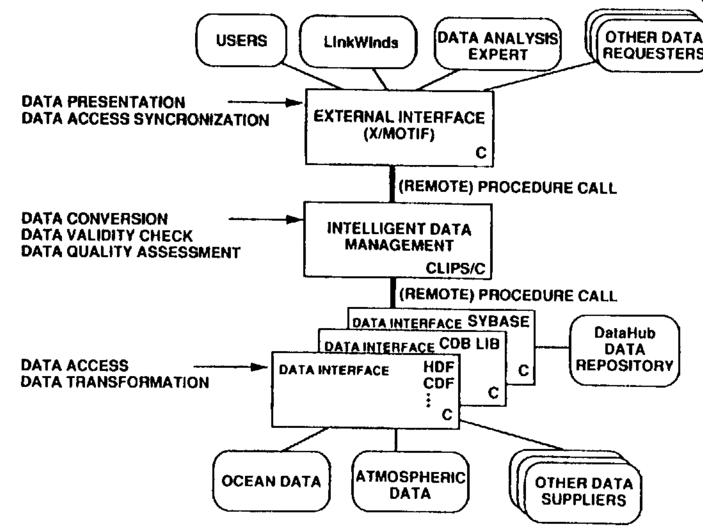
TRANSACTION:	<TRANSACTION-NAME>
IN:	(<OBJECTLIST>)
OUT:	(<OBJECTLIST>)
PRECONDITION:	<PREDICATES>
ACTIONS:	<TRANSACTIONS AND/OR OPERATORS>



ASPRS/ACSM/RT 92 CONVENTION

JPL

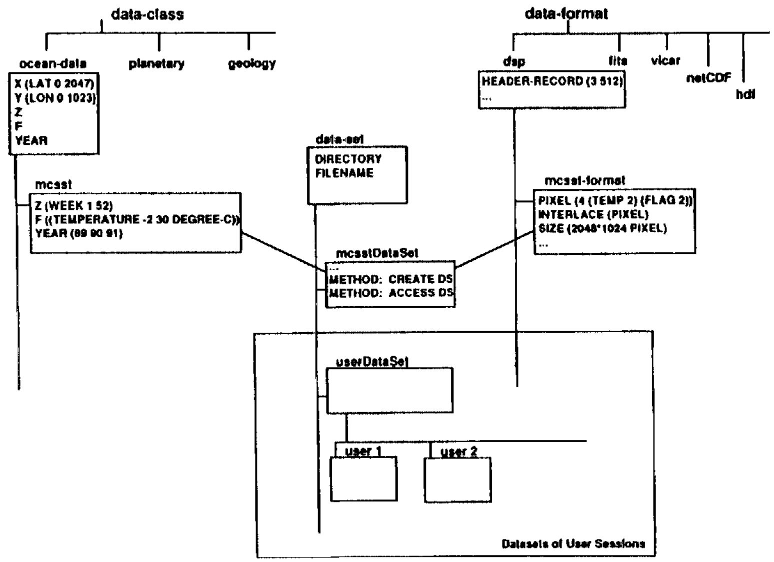
SOFTWARE ARCHITECTURE



ASPRS/ACSM/RT 92 CONVENTION

JPL

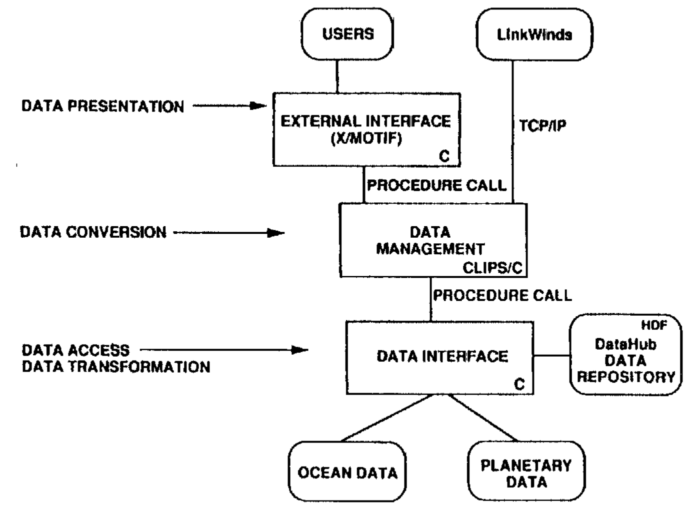
USER CREATED DATASETS



ASPRS/ACSM/RT 92 CONVENTION

JPL

SOFTWARE IMPLEMENTATION



ASPRS/ACSM/RT 92 CONVENTION

JPL

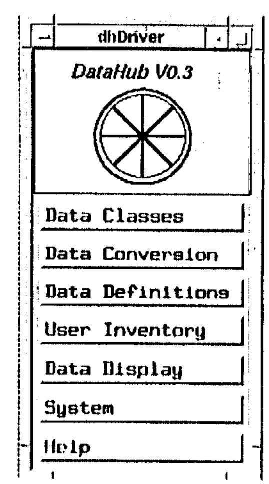
STATUS

- DEFINED A GENERAL FRAME WORK FOR SCIENCE DATA MANAGEMENT
- IDENTIFIED A CRITICAL SUBSET OF DATA OPERATORS FOR THE SCIENCE DATA VISUALIZATION APPLICATIONS
- INITIAL PIOTOTYPE-DataHub 0.3-
  - COMMON USER INTERACTION
  - DATA FORMAT CONVERSIONS
    - DSP: MCSST, CZCS
    - VICAR-LABELED: AVIRIS, VOYAGER 1&2
    - PDS-LABELED: MAGELLAN, MDIM
    - CDF: IN-DEVELOPMENT
  - USER INVENTORY MANAGEMENT
  - DataHub/LinkWinds INTERFACE
  - UNDERLYING OBJECT-ORIENTED STRUCTURES AND METHODS

ASPRS/ACSM/RT 92 CONVENTION

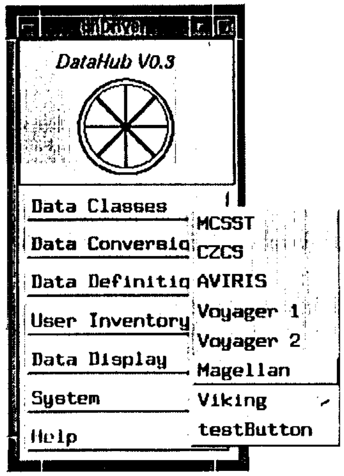
JPL

MAIN MENU



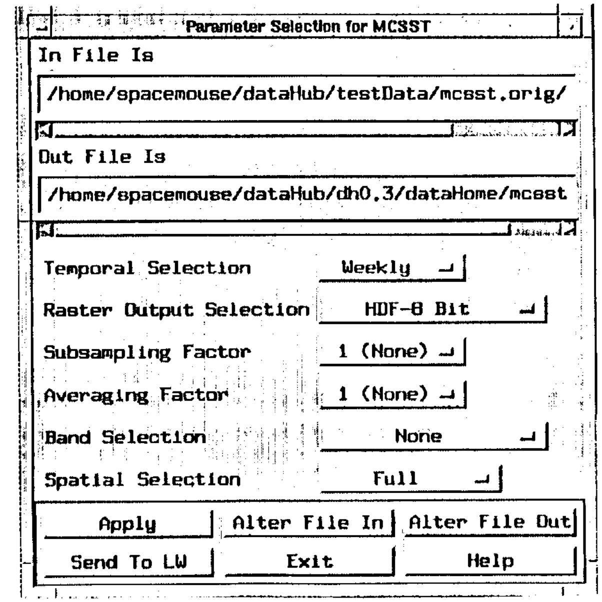
ASPRS/ACSM/RT 92 CONVENTION

DATA CLASSES



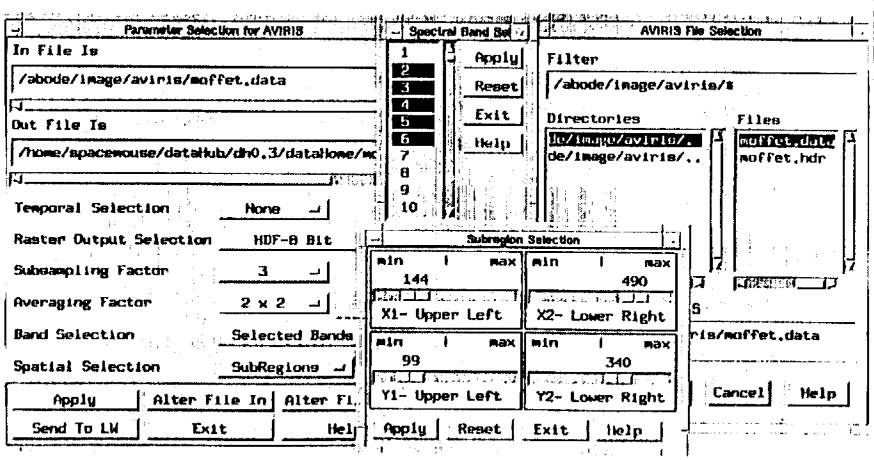
ASPRS/ACSM/RT 92 CONVENTION

CENTRAL SELECTION CHOICES



ASPRS/ACSM/RT 92 CONVENTION

AVIRIS SELECTIONS



ASPRS/ACSM/RT 92 CONVENTION

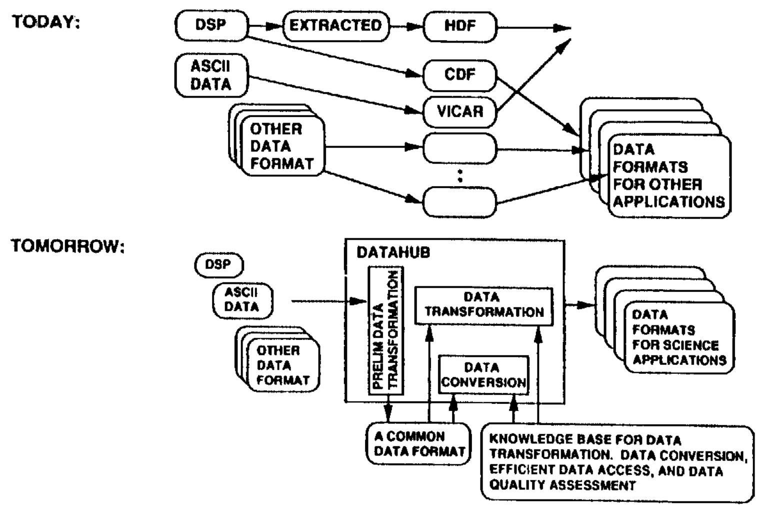
FUTURE WORK

- USING THIS FIRST PROTOTYPE VALIDATE BASIC CONCEPTS WITH THE END-USERS
- ADDITIONAL DATA SETS
  - USER-DEFINED DATA FORMATS AND CONVERSIONS
- EXPAND KNOWLEDGE ABOUT THE DATA
  - DATA FORMATS – USAGE OF METADATA IN HEADERS
  - DATA SEMANTICS – MEANING OF DATA VALUES, RELATIONSHIPS BETWEEN DATA SETS, DISCIPLINE DEPENDENT DATA ACCESS/ANALYSIS METHODS
  - DATA SEMANTICS METHODS AS REPRESENTED BY USER'S CONTEXT IN THE VISUALIZATION REGIME
- EXPAND/ENHANCE OBJECT ORIENTATION
- USER INVENTORY
- ADDITIONAL UNDERSTANDING OF ANALYTICAL TOOLS (PREPROCESSING SUPPORT)
- DATA PRESENTATION – EXCHANGE PROTOCOLS THAT FACILITATE EXPLORATORY DATA ANALYSIS

JPL

ASPRS/ACSM/RT 92 CONVENTION

SCIENCE DATA MANAGEMENT



JPL

ASPRS/ACSM/RT 92 CONVENTION

LinkWinds
OBJECTIVES

- DEVELOP A SOFTWARE ENVIRONMENT TO SUPPORT THE RAPID PROTOTYPING AND EXECUTION OF DATA ANALYSIS/VISUALIZATION APPLICATIONS
- PROVIDE A SUITE OF TOOLS TO INTERACTIVELY VISUALIZE, EXPLORE AND ANALYZE LARGE MULTIVARIATE AND MULTIDISCIPLINARY DATA SETS
- DEVELOP A USER INTERFACE WHICH ALLOWS MAXIMUM DATA AND TOOLS ACCESSIBILITY WITH A MINIMUM OF TRAINING
- PROVIDE TOOLS TO MAKE THE ENVIRONMENT ACCESSIBLE TO APPLICATION DEVELOPMENT BY USERS

JPL

ASPRS/ACSM/RT 92 CONVENTION

LinkWinds
DATABASE INTERFACE

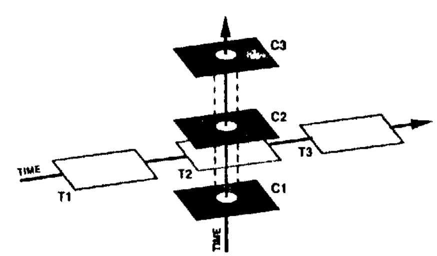
- CURRENTLY STANDARDIZED ON HIERARCHICAL DATA FORMAT (HDF) CREATED AT UNIVERSITY OF ILLINOIS, CHAMPAGNE/URBANA, NCSA
- OTHER DATA FORMATS ARE ACCESSIBLE VIA INTERACTION WITH DATA HUB
- USER INTEGRATES DATABASE FILES INTO SYSTEM BY EDITING THREE TEXT FILES CONTAINING DATA DESCRIPTION AND METADATA. TEMPLATES ARE PROVIDED

JPL

ASPRS/ACSM/RT 92 CONVENTION

RESULTS FROM ARTIFICIAL NEURAL-NET MODEL

- DEVELOPED ALGORITHM TO MODEL/INTERPOLATED COASTAL ZONE COLOR SCANNER (CZCS) TO REPRESENT PIGMENT CONCENTRATION
- USED BOTH TEMPERATURE AND PIGMENT CONCENTRATION TO FORM AN ESTIMATE OF THE MISSING DATA IN THE PIGMENTS FIELD
- BOTH THE SPATIAL AND TEMPORAL COVERAGE OF TEMPERATURE AND PIGMENTS ARE USED AS INPUT PARAMETERS FOR THE ARTIFICIAL NEURAL-NET MODEL



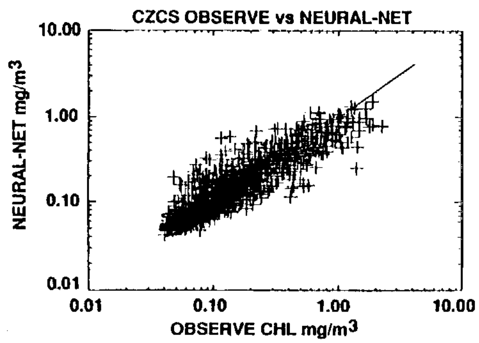
- THIS REPRESENTS HOW SPATIAL AND TEMPORAL INFORMATION ARE USED IN THE ARTIFICIAL NEURAL-NET; WHERE C AND T REPRESENT PIGMENT CONCENTRATION AND TEMPERATURE RESPECTIVELY

ASPRS/ACSM/RT 92 CONVENTION

JPL

VERIFICATION OF THE TECHNIQUE

- 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 HAS BEEN MODEL BY THE ARTIFICIAL NEURAL-NET



- CORRELATION BETWEEN PIGMENT CONCENTRATION FROM SATELLITE MEASUREMENTS AND RESULT FROM NEURAL-NET IS $R_2 = 0.92$

THE STATE OF SCIENTIFIC VISUALIZATION WITH REGARD TO THE NASA EOS MISSION TO PLANET EARTH

Dr. Michael E. Botts
University of Alabama in Huntsville

August 12, 1992

THE STATE OF SCIENTIFIC VISUALIZATION WITH REGARD TO THE NASA EOS MISSION TO PLANET EARTH

A Report to NASA Headquarters

August 5, 1992

[Final Draft]

Michael E. Botts, Ph.D.
Earth Systems Science Laboratory
University of Alabama in Huntsville (UAH)
Huntsville, AL 35899

Executive Summary

THE IMPORTANCE OF VISUALIZATION TO THE EOS MISSION

- Visualization is a vital component of the tools required for meeting EOS scientific objectives. Neglecting visualization could result in failure to meet these objectives.
- At present, NASA has given considerable attention to the data retrieval, data management, and scientific objectives nodes of the data pipeline, but has given minimal attention to the role of visualization within the EOS project. The justification for this approach has been primarily rooted in the belief that Commercial Off-The-Shelf (COTS) software, as well as modest development activities of NASA and other agencies, will be adequate to meet these needs.
- This initial report highlights successes and general deficiencies with the current state of visualization development and application. It is based on the results of meetings with 50 groups of EOS scientists and developers at 30 sites, as well as the author's experiences in trying to meet visualization needs at NASA MSFC. A follow-up report will examine possible options for assuring that there are adequate and proper visualization tools for meeting the scientific objectives of the EOS mission.

APPLICATION OF VISUALIZATION WITHIN THE EOS MISSION

- Applications of visualization within the EOS project include:
  - Scientific Investigation
  - Data Validation
  - Model and Algorithm Development and Validation
  - Data Browse
  - Information Transfer
  - Mission Operations

BACKGROUND ON VISUALIZATION

- Visualization, although a relatively new discipline, actually consist of the unification of several fairly mature components, including:
  - Image Processing
  - 2D Data Plotting
  - 3D Computer Graphics and Animation
  - Volume Rendering
  - Geographic Information Systems (GIS)
  - Computer-Human Interactions (CHI)

Each component provides important capabilities to Earth systems scientists, but, at present, each also has specific deficiencies for meeting EOS needs.

CHANGING ROLES AND CAPABILITIES

- Due to rapidly increasing CPU and graphics power available to the scientist at his desktop, two major transitions are occurring within the scientific computing environment, including transitions:
  - from centralized to distributed computing
  - from batch-mode operations to interactive computing

  Both of these transitions are putting more computing power and control directly in the hands of the scientist. It is this interactivity that will result in the greatest benefits to be derived from scientific visualization. However, without the proper software to take advantage of this power, these benefits will not be realized.
- These transitions are demanding changes in the roles played by scientists and computer specialists, as well as requiring changes to our visualization and analysis tools. The batch mode of handing off visualization jobs to computer specialists is no longer adequate.
- The movie making era of visualization was a necessary and important phase of these transitions, but does not represent the total required direction for visualization.
- The most important immediate direction for visualization efforts is that of putting useful and usable interactive tools into the hands of the scientist.

WHY AREN'T SCIENTISTS USING WHAT'S AVAILABLE?

- Many of the techniques and components of visualization required for meeting EOS scientific objectives are available today. However, the actual use by the scientist of even our present visualization capabilities is well behind these capabilities.
- Why aren't scientists using the visualization capabilities that are available to them today? Reasons include:
  - The tool is not extensible or is too inflexible.
  - The tool is too difficult to learn and use.
  - It is too difficult to get existing data into the tool.
  - The tool does not adequately link visualization and analysis.
  - The collection of tools, as well as the data, exist within a complex heterogeneous computing environment.
  - The tool does not do what the scientist needs to do.
  - The scientist is not aware that the tool exist or that it meets his/her needs.
  - The tool is too costly.

– The scientist does not have access to adequate hardware for running the tool.
– It is too difficult to communicate the results of the finding to others, because (a) it is too difficult to get color prints or video, (b) it is difficult or impossible to interact remotely with colleagues, or (c) the publishing industry is too technically and philosophically archaic to meet the present needs for color hardcopy, animation videos, algorithm and application exchange, and voice and sound annotation.

WHERE ARE THE VISUALIZATION BOTTLENECKS ?

- With few exceptions, hardware capabilities are not, at present, a major limitation of our visualization environment. The commercial marketplace is probably adequate for assuring necessary advances in hardware technology.

- The primary bottleneck is the lack of adequate software which allows the scientist to take advantage of this power and to interactively visualize and analyze his/her data within our complex computing environment. It is questionable whether Commercial Off-The-Shelf (COTS) software will be adequate for meeting all of the visualization needs of EOS.

COTS, PUBLIC DOMAIN, AND IN-HOUSE DEVELOPMENT

- There are advantages and disadvantages to relying on either COTS or in-house developed software for meeting the visualization needs of EOS. A proper balance between COTS, public domain, and in-house development is advantageous, but this balance must be accomplished with adequate and properly-directed support.

- If NASA is to rely more heavily on COTS software, then it must re-evaluate and improve the way that it deals with COTS developers.

CURRENT DEVELOPMENT ENVIRONMENT WITHIN NASA

- NASA funded development has resulted in several significant, leading edge visualization and analysis tools.

- Unfortunately, with particular regard to visualization, the development environment within the OSSA can be characterized as:

  – Fragmented, with little overall direction or coordination
  – Lacking an established and understandable organizational structure for funding both general tool and application-specific development
  – Lacking a firm commitment or plan for meeting the true visualization needs of the EOS science community
  – Lacking adequate mechanisms for technology transfer both within OSSA and between OSSA and OAST
  – Experiencing inefficient use of limited funding

- Required in-house development is of two types: (1) development of general tools to meet the needs of a wide range of users, not being met by existing software, and (2) extension, modification, and integration of existing tools to meet application-specific requirements.

DIAMONDS AND DINOSAURS

- A difficult challenge for NASA will be that of recognizing and supporting in-house "diamonds" (experimental development efforts which hold much promise for scientists) and of preventing, recognizing, and dealing with "dinosaurs" (old development programs with limited momentum and decreasing application). Present mechanisms for transferring successful development activities into usable technology is inadequate.

MAIN AREAS FOR CONCENTRATION

- The main general areas presently needing consideration include:

  – **Integration of Visualization with Data Management and Analysis** - Data management, analysis, and visualization represent a triad of functionality required to meet scientific objectives of the EOS mission. To meet these objectives, this triad must be properly balanced and adequate bidirectional links must exist between each component. Links from visualization to analysis, and from visualization to data management, are essentially nonexistent, at present. Navigation, or knowing the spatial and temporal location of each data point, spans all components of this triad and is a particularly critical subset of this issue requiring immediate attention.

  – **Application Development for the Scientist as the End-User** - The most important immediate need for visualization is that of getting interactive tools into the hands of the scientist. This requires that these tools meet the actual needs of the scientist, be simple and intuitive to use, and be logical to the scientist rather than to a computer specialist.

  – **Providing Extensibility without Complexity** - Extensibility of visualization tools is a vital requirement, particularly for COTS software. This extensibility must be provided without a significant increase in complexity of use, or any resulting complexity must be able to be hidden from the scientist by customizable interfaces.

  – **Application-Specific Programming without Redundant Programming** - Redundant programming within the Earth systems science community is extensive. Often new development efforts are undertaken in order to provide functionality not available from existing software. However, in order to provide new application-specific functionality, many development activities spend a very large portion of its efforts on redundantly programming functionality that already existed in several other programs. The availability and application of extensible software and the adoption of object-oriented programming principles would minimize redundant programming while meeting application-specific needs.

  – **Simplification of the Complex Heterogeneous Computing Environment** - Scientists and computer specialists are forced to operate within a complex heterogeneous computing environment, consisting of incompatible operating systems, graphics protocols, networks, file formats, and output devices. Efforts to homogenize this environment will help to a limited degree, but would probably be only temporary relief. There is a vital need for utilities which shield scientists from dealing with unnecessary complexities, allowing them to concentrate on analysis of data.

  – **Development of New Techniques and Components** - The exploration of new techniques for analyzing and visualizing data should continue. In a few cases, very different approaches to analyzing and visualizing data may be introduced which require development of a complete application. However, in most cases, the introduction of new techniques should require only the development of a module or component which can be added to existing applications.

  – **Education and Communication** - Many of the present challenges in visualization are the result of a lack of proper communication and education. Scientists and project managers must become better educated as to the scientific benefits of visualization and the availability of existing software, while in-house and COTS developers must become more aware of the true needs and objectives of the scientist.

  – **Distribution and Maintenance** - The distribution, maintenance, and support of COTS, public domain, and in-house software is a significant challenge. Proper mechanisms for licensing and supporting COTS software within our distributed environment should be in place to assure their availability and testing. A major challenge with in-house development is the distribution and maintenance of successful application development, and the transition of such programs from the experimental stage through maturing and operational stages.

CONCLUSIONS

- Visualization is vital for meeting the scientific objectives of the EOS mission.

- Although our present suite of visualization techniques is impressive and powerful, the application software for putting these capabilities into the hands of the scientist is, at present, inadequate.

- The development of extensible, user-friendly, object-oriented software within the commercial software industry, is helping to increase the probability that COTS software can serve as a core for meeting many visualization needs. However, there will still be a need for in-house efforts directed at extending and modifying these tools to meet application-specific requirements, and for developing leading edge techniques and applications not available from the commercial front.

- A follow-up report will investigate viable options for meeting the visualization requirements of the EOS mission in the future. These options will focus on two major objectives: increasing the effective use of visualization tools by scientists, and maximizing the return on development efforts.

Table of Contents

THE IMPORTANCE OF VISUALIZATION TO THE EOS MISSION

The proper storage, management, and distribution of the enormous database to be generated by the NASA Earth Observing System (EOS) Mission is extremely important to the scientist requiring access to this data. However, without the proper tools to also analyze these data sets, this enormous database is strictly that, a database. The successful accomplishment of the scientific objectives of EOS is highly dependent on the scientists' ability to analyze and interact with these data in order to retrieve the important knowledge hidden within this database. The only way to effectively retrieve significant information from such an enormous database is through a well-balanced use of data management, analysis, and visualization.

As illustrated in the simplistic cartoon of Figure 1, the typical data flow of EOS-related data moves from the satellite and ground-based data sources, into the data storage and data management nodes, and to the analysis/visualization nodes where the scientist gains scientific insight through his perception of the analytical and visual results. Results must then be communicated to colleagues or policy makers through speech or visual means. A bottleneck at any point along this path flow will result in less than successful completion of the scientific objectives of EOS. At present, NASA has given considerable attention to the data retrieval, data management, and scientific objectives of the data pipeline, but has given minimal attention to the important role of visualization. The justification for this approach has been primarily rooted in the belief that Commercial Off-The-Shelf (COTS) software, as well as modest development activities of NASA and other agencies, will be adequate to meet these needs.

The purpose of this study is to investigate the needs and application of visualization tools relative to the EOS objectives, to become familiar with mature and on-going visualization development efforts of industry, NASA, and other groups, and to evaluate whether these efforts will be adequate for meeting the scientific objectives of EOS. The results are based on the results of meetings with 50 groups of EOS scientists and developers at 30 sites, as well as the author's experiences in meeting visualization needs at the NASA Marshall Space Flight Center (MSFC). This initial report highlights successes and deficiencies with the current state of visualization development and application, and recognizes areas where attention should be focused. A follow-up report will examine possible options for assuring that we have adequate and proper visualization tools for meeting the scientific objectives of the EOS mission.

APPLICATION OF VISUALIZATION WITHIN THE EOS MISSION

There are five major applications of visualization within the EOS mission:

- **Scientific Investigation**
- **Data Validation**
- **Model/Algorithm Development and Validation**
- **Database Browse**
- **Information Transfer**
- **Mission Operations**

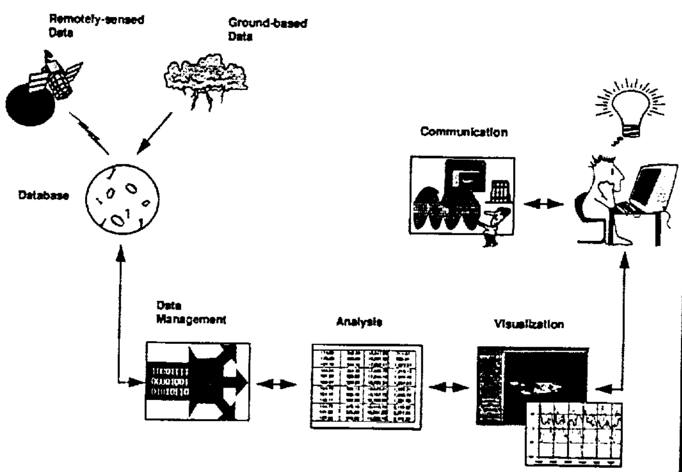


Figure 1. A cartoon illustrating the typical data flow of EOS related data from ingest to scientific insight.

[*** Incomplete Figure ***]

Visualization is proving invaluable in providing new insight into Earth systems science problems. Whereas scientists were previously limited to painstakingly analyzing a small subset, or profile, of the an available data set, visualization is now providing scientists with the ability to observe and analyze complete data sets within 4D spatial and temporal domains. Even more significant is the ability to view, analyze, and correlate several data sets, related in space and time, in order to better discern the relationships and interactions of various physical factors. The interactivity of these tools allows the scientist to more effectively probe the data and the analysis process, and to receive immediate response to "what-if" questions.

Even simple animations have been shown to be very useful in pinpointing errors within data and computer models, and in many cases, in tracing the source of these errors. This is particularly important when locating the source of propagating errors within model runs. Such errors may not be spotted, otherwise, unless they result in fatal errors, and even then it becomes quite difficult to trace back to the source of the error without the assistance of visualization. As visualization techniques become better combined with analytical methods, data and model validation, using visualization, will become much more effective and efficient.

The abundance and size of data sets available to an EOS investigator, will necessitate the use of data browse capabilities. The efficient use of network traffic, computer resources, and the scientist's time will suffer significantly without the scientist's ability to quickly transfer and visualize greatly compressed data subsets, in order to predetermine the scientific relevance of the full data set.

Finally, visualization offers one of the most effective and intuitive means of communicating large amounts of scientific information to colleagues, program managers, policy makers, and students. This importance of visualization for education and communication should not overlooked. Furthermore, as remote interaction of science via graphics workstations, becomes more a reality, the use of visualization for information exchange between scientists will become routine and very effective.

BACKGROUND ON VISUALIZATION

Although the field of scientific visualization is still in its infancy, it actually involves the unification of several traditionally independent disciplines, some of which are fairly well developed. These include:

- **2D Data Plotting and Contouring**
- **Image Processing**
- **3D Computer Graphics, CAD, and Computer Animation**
- **Volume Rendering**
- **Geographical Information Systems (GIS)**
- **Computer-Human Interaction (CHI)**

Each of these disciplines has provided, and will continue to, provide significant tools for scientific investigations. Each also has specific areas of deficiencies which will be discussed in more detail in a later report. For example, many of the 2D and 3D plotting applications still retain the mentality of pen-plotter output, which does not allow interactivity between the user and the plot displayed on a screen. Many visualization application programs have completely ignored the power and simplicity of plots and contour maps for displaying important information, much to the dismay of the science community.

Although digital image processing is a very important component of visualization, many techniques for processing imagery are rather archaic, particularly with regard to fully utilizing data available in more than three spectral bands. Also, many of the major application programs for image processing have also become somewhat archaic with regard to their user interfaces, program architectures, graphics capabilities, data structures, interactivity, and portability to RISC-based workstations.

Advances in 3D computer graphics, CAD, and computer animation, made during the 1980's, provide a substantial base for the development of important tools for visualizing 3D, temporal, scientific data. However, having been developed primarily for the purposes of military simulation and broadcast animation, many of the algorithms and application programs don't meet the needs of the visualization community today. Similarly, while volume rendering holds much promise for detailed visualization of volumetric scientific data, most of the present algorithms were developed for medical applications and are not well suited for interactivity, or for dealing with temporal or non-Cartesian data.

Likewise, GIS developed primarily within cartographic-based, rather than scientific, disciplines. As a result, several present deficiencies in traditional GIS techniques have been recognized which limit its effectiveness in scientific applications. These include the lack of effective means for dealing with temporal elements, the inability to recognize and measure potential errors resulting from different spatial and temporal resolutions between data sets, and the inadequacies for handling of "fuzzy", or continuous, data boundaries.

Finally, the field of CHI is providing important insight into effective interaction between man and computer. The design and application of proper graphical user interfaces (GUIs) and intuitive program structure are crucial to putting effective visualization tools into hands of the scientist. In addition, a better understanding of factors which affect human perception is certainly important to the design of visualization tools, but is an area that is often overlooked during their development.

Whereas the '80s was the decade for rapid advancement of computer graphics related to broadcast animation, military simulation, and industrial CAD, the '90s is promising to be the decade for scientific visualization. Many of the present efforts have simply involved the application of individual components, such as computer graphics and animation, to scientific data. Certainly, the techniques developed within each of these components represent a substantial base for visualization capabilities. However, the success of visualization as a scientific tool will depend on the proper adaptation of these components to meeting the needs of the science community, as well as seamless integration of all of these components into application programs which are useful and usable to the scientist.

It is not only important that we recognize the deficiencies that exist in our present tools, but that we also understand the current and future environment under which improvements to our tools must be made. These are examined in the following section.

CHANGING ROLES AND CAPABILITIES

The CPU and graphics power available to the scientist is increasing at a logarithmic rate of an order of magnitude every 3-5 years. At the same time, this power is becoming available on low-priced workstations. This reality is moving visualization computing away from a highly centralized environment, in which expensive computer equipment was centrally located and only highly specialized personnel operated this equipment, to a more distributed environment, where the scientist has significant computing power at his desk, while still having access to greater computing power and large amounts of data over the network. The scientist is thus becoming more in control of his computing environment, as well as taking on greater responsibility for the operation of his computing tools.

In addition, the increase in computing power and the decentralization of the computing environment are allowing much greater interactivity than was previously possible. Assuming that the proper visualization and analysis tools are in place, it is this interactivity between the scientist and his/her data that will result in significant advances in the Earth systems science community. The batch mode of computer operation is no longer adequate for meeting the demands of scientific analysis.

We are presently in a state of transition from centralized to distributed computing, and from batch to interactive processing. These transitions are demanding changes in the roles played by the scientists and computer specialist, as well as requiring changes to our visualization and analysis tools.

As an example of these changes, consider the roles and environments existing in typical visualization activities of the past and present. Previously, it was adequate for a scientist to request, from an image processing specialist, that he or she apply a principle component analysis to a given Landsat scene, stretch the resulting bands over 8 bits, create an RGB composite of principle components 2, 3, and 4, and provide a color print of the resulting image. The required batch operation was easily defined by the scientist, who could in the end study the final image for weeks or months, creating hand-drawn overlays to illustrate scientific findings.

In contrast, the scientist of today might wish to study and correlate several data sets which are related in 3D space and time, including, for example, satellite imagery, data measured along a flight path, 3D volumes from radar, and 3D gridded data from a numerical model. In order to gain significant insight from this study, the scientist will probably require, at a minimum, the ability to render representations of these data in 3D space, interactively move the point of view around the 3D space, view and control the animation of the data through time, control the representation of the data (e.g. color maps, transparency, etc.), evaluate the data with an appropriately-placed 3D probe or 2D slice plane, and generate relevant plots "on-the-fly".

There is no longer a definable product that can be demanded of the computer specialist by the scientist, other than that of providing the scientist with proper tools, usable by the scientist, of properly integrating the scientist's data into the tool, and of providing operational training and assistance when required.

During this transition period, it is expected that we should see some remnants of our previous roles and computing environments. In fact, the proliferation of sophisticated movies showing scientific visualization is somewhat a symptom of this. While these movies have undeniably provided benefits in the form of scientific insight, public relations, and awareness of the power of visualization, they are not the future direction for visualization that is required to meet the total needs of the scientist. Far from being interactive nor analytical, these movies have often been created at great cost in time and funding, typically requiring weeks to months of effort from teams of scientists, programmers, computer visualization experts, computer animators and artists, and multimedia specialists. While this movie making is a natural extension of the 80's style of computer animation, as well as an extension of the previous roles of the computer specialists, many short-sighted scientists have, as a result, developed the view that visualization is strictly for creating "pretty pictures" and not a tool for doing significant science. In others, these movies have instilled false security that the required visualization tools required are in place. In reality, scientific animation and movie making represent an important part of visualization, but they are only a small part of the requirements for visualization.

In essence, the most important immediate direction for visualization efforts is that of putting these tools in the hands of the scientist, allowing him/her to interactively probe the data without the constant aid of computer specialists. Computer visualization specialists should be involved in getting these tools into the scientist's hands and modifying and expanding these tools as required, and not in operating the tools for the scientist. Thus, the role of the computer specialist is moving away from the backend user of the tool, to the frontend gatherer, tester, and modifier of the tool. This, of course, requires that the application be well designed for the scientist as an interactive end user, with particular attention given to meeting the scientist's needs in a simple and intuitive manner.

In view of this, it is important to recognize the diversity that exists among scientist with regard to their application of computer tools. While not based on any statistical surveys, it appears that approximately 20% of the scientists are comfortable with and actively involved in finding, learning, modifying, and applying the computer tools that are available, and in many cases are willing to work with developers on improving these tools. At the other extreme, are approximately 20% of the scientists, who believe that the tools that they have employed in the past are adequate for meeting their needs, or who, because of lack of time or desire, only wish to be given a tool that requires very little time or attention to learn and operate. In lieu of such a tool, this group of scientists will rely heavily on computer specialists for all aspects of tool application. The remaining 60% of the scientists are reluctant or unable to concentrate a lot of attention on visualization and analysis tools, but will invest the required time to test and learn new tools if properly nurtured and if benefits become apparent.

Scientists who are willing to invest time and effort in the testing and modification of tools are a very vital asset to both the developers of these tools and to the scientific community, as a whole. These scientists should be encouraged and supported, not ostracized by other scientists and funding agencies as has occurred in the past. However, it is also important for developers to be aware that these particular scientists represents a select group and not the majority of scientists who will not use a tool that is difficult to learn or use, or that does not immediately meet their needs.

WHY AREN'T SCIENTISTS USING WHAT'S AVAILABLE?

With a few exceptions to be discussed in a later report, many of the techniques and components of visualization required for the EOS mission are available today; these will probably progress and others will continue to be developed regardless of NASA's direct involvement. Certainly, NASA's support of new techniques for visualization and analysis would greatly improve our future abilities and would direct development efforts along paths that are critical to NASA's objectives.

In essence, although many of the components do exist to meet many of our present and future visualization requirements, the actual use by the scientist of even our present visualization capabilities is well behind the development and availability of these capabilities. The major challenges and deficiencies with visualization, at present, lie in the integration of these components and techniques into applications that are usable and useful to the scientist in his everyday work.

A major question to ask, then, is "WHY AREN'T SCIENTISTS USING THE VISUALIZATION CAPABILITIES THAT ARE AVAILABLE TO THEM TODAY?" The answers to this question begin to provide an understanding of the directions required for improving the state of visualization within the scope of the EOS mission. Although not every problem discussed below applies to every available application program, most every visualization tool or collection of tools required for a particular task suffers from one or more of the following deficiencies:

- **The tool is not extensible and is too inflexible** to allow scientist to add discipline specific functionality or incorporate existing modules; thus, too many tools are often required to meet the scientist's tasks.

- **The tool is too difficult to learn or use.** This is particularly troublesome if more than one tool is required to meet scientist's needs. Difficulties of use can exist due to archaic command line interfaces or as a result of program structures and graphical user interfaces that are intuitive to computer experts, but not scientists.

- **It is too difficult to get existing data into the tool**: this is generally a result of (a) incompatible data formats and the lack of available data filters, (b) the lack of true integration between the visualization/analysis tool and the data management system, and (c) the exposure of the scientist to the complexities of networking.

- **The tool does not adequately link visualization to analysis**, which is critical to the scientific process. Going "back to the numbers" or more importantly, "back and forth between the visual and analytical" is crucial before visualization can become a "real" science tool in the scientist's mind.

- **The collection of tools, as well as data, that are required for meeting a particular science need, exist within a complex heterogeneous computing environment.** For example, the required tools might be running on different architectures with different operating systems, data might exist on different hosts as incompatible binary files, and the application might require different data formats. In addition, output of results may require completely different formats, complicated commands, and the need to transfer data to some other device on the network.

- **The tool does not do what the scientist needs to do**, either because no tool can meet all needs or because the tool was designed with computer graphics, rather than science, in mind.

- **The scientist is not aware that the tool exist or lacks appreciation of the visualization tool** as a serious tool that could greatly benefit his or her scientific research, or model and data validation.

- **The tool is too costly**, particularly for desktop visualization.

- **It is too difficult to communicate the results of the findings to others.** This results from at least three causes: (a) it is too difficult to get prints or video from the tool or platform, (b) it is impossible or difficult for the scientist to remotely interact and discuss with other scientist, the intermediate and final visual results derived from the tool, and (c) the publishing world is not well equipped, nor philosophically structured, to deal with the present needs for color hardcopy, animation videos, algorithm and application exchange, and voice and sound annotation.

- **The scientist does not have access to hardware required** to run the application software, or the scientist's hardware is too graphically underpowered to run the application effectively.

Because of these difficulties, the application of present visualization tools have been inadequate and frustrating to scientists and computer specialists, alike. Many are not able to dedicate the time required to find and learn visualization tools which might meet their requirements, and have become frustrated and disenchanted with the present tools that are available to them. In order to "get the job done", many scientists often give up using present visualization tools in lieu of more familiar, but inadequate, analysis methods. The scientists who do persevere often reap significant benefits from the present visualization tools. However, a significant amount of their time and concentration, which could be applied to gaining insight into the data, is often lost to learning, operating, and modifying the present tools.

WHERE ARE THE VISUALIZATION BOTTLENECKS ?

With a few exceptions, hardware capabilities are not, at present, a major limitation of our visualization environment. The present hardware capabilities available in the market, greatly exceed the present availability of software tools with which the scientist can routinely take advantage of this power. Certainly, our increasing demands for very large data sets, our increased network traffic resulting from more distributed processing, the transfer of large files, increased multimedia activity, and the use of X window protocol, and our need for interactivity in visualization and analysis, are resulting in a constant requirement for higher CPU and graphics power, and wider bandwidths throughout the data path. However, with the possible exception of network requirements, the commercial marketplace is probably adequately suited for assuring that the capabilities of available hardware will continue to increase at a rate commensurate with our demands and with our ability to take advantage of the available power.

With the capabilities of computers increasing by an order of magnitude every 3 to 5 years, it is important to consider the short time frame in which our capabilities increase and our present computers become obsolete. In addition, considering that the computer is possibly the most vital tool of today's scientist, it is important that funding and expedient procurement of these tools be available. The lack of accessibility to adequate hardware power, particularly at universities, and lack of expedient procurement of required hardware, particularly at government institutes, are at present limiting the effective use and development of visualization tools in the scientific community.

Still, the primary bottleneck in visualization, at present, is the lack of adequate software which would allow the scientist to interactively visualize and analyze his/her data within a complex heterogeneous environment, without the need for the scientist to be concerned with the underlying complexities of the tools. Although there have been some significant advancements in visualization software, the available software as a whole is, at present, deficient in meeting these needs. Unlike hardware, it is questionable whether Commercial Off-The-Shelf (COTS) or public domain software will be adequate for effectively meeting the total visualization requirements for EOS. The following section examines this issue in more detail.

COTS, PUBLIC DOMAIN, AND IN-HOUSE DEVELOPMENT

The question arises as to what extent the needs for scientific visualization software can be met by relying on the commercial software industry or other outside sources, in lieu of development funded by NASA. Each source has advantages and disadvantages.

The advantages of relying on COTS include:

- NASA pays only for working product not development or experimental failures
- Products tend to be more polished
- Maintenance and support not as major an issue for NASA
- There is generally more awareness of a product's existence

The disadvantages of relying on COTS include:

- Applications have traditionally been inflexible and not extensible
- There is no access to source code required for code modification or porting of the application to necessary platforms
- Applications tend to be designed for large market appeal, not application specific needs
- There is a higher cost factor to user, but not necessarily to NASA as a whole
- Users are less apt to "try it out" if the software must be purchased first
- There is no direct control of directions of future development, and no assurance that developers will be around next year
- Application is often developed outside of the scientists' realm, generally without a true understanding of scientific needs
- Approach for meeting needs is generally more conservative and less creative

In contrast, the advantages of relying on in-house development are:

- Software is designed to meet general and specific needs of NASA's scientific objectives
- Source code is available to all for modification and portability
- The product is generally flexible and adaptable, or can be cannibalized for new applications
- New creative and leading edge products often result

The disadvantages of relying on in-house include:

- NASA assumes the risk of development with no assurance of success
- Maintenance, distribution, and support are significant challenges
- There is a risk of the software becoming a costly dinosaur that NASA is forced to maintain
- Without proper direction and coordination, there is traditionally much redundancy of programming with in-house development

There are many commercially available visualization packages which are well ahead of most in-house developed software. Although some of this commercial software arose from initial development efforts at NASA and other governmental institutes, they have generally become more polished and portable. There are also many NASA-funded development efforts that are unmatched in functionality by the software industry. Still, in many cases, COTS software can probably provide the core for many visualization needs, particularly those products which are more extensible. However, if NASA intends to rely significantly on COTS software, then consideration should be given to changing the level of interaction with vendors. In particular, NASA should encourage more communication between scientists and the developers of COTS software.

In addition, there are many significant software contributions available within the public domain. Although a few of these, such as Khoros and MATLAB, are large, refined development efforts, most are primarily small-scale efforts which provide useful utilities or limited applications. Since the source code is generally available, public domain software is a good source for useful utilities or usable algorithms. Since maintenance and support with these packages is often a challenge for the developers, particularly for larger-scale efforts, NASA should consider means of providing assistance for packages which it feels provide significant benefits to EOS scientists.

While COTS software may play a significant role in meeting EOS visualization needs, relying solely on development from the general software industry may not be adequate. Particularly without guidelines from NASA, and changes to the way that NASA deals with vendors, it is risky to assume that the outside industry will take on the challenge to develop visualization software capable of meeting the needs of the EOS program. This is particularly true for application-specific needs which may not provide an adequate commercial market.

A balanced approach between COTS and in-house development is probably the best solution for meeting the visualization requirements of the EOS mission. The needs for in-house visualization development within NASA include:

- Extension or modification of existing software to meet application-specific requirements
- Contributions of useful modules or components to the general programming pool
- Leading edge (experimental) development of techniques and applications
- Application-specific programs not available from COTS and not able to be created by simple extension of existing programs

These needs should be recognized and programs should be structured to adequately provide support for meeting these needs.

CURRENT DEVELOPMENT ENVIRONMENT WITHIN NASA

Development activities throughout NASA have resulted in some significant, leading-edge visualization and analysis tools. Both advanced and experimental visualization development has transpired within the Office of Aeronautics and Space Technology (OAST) to support flight operations, aeronautical development, computational fluid dynamics (CFD) research, and virtual reality. Although not presently well suited for meeting visualization needs of EOS, some of these efforts could be modified to meet some of these requirements.

Within the Office of Space Science and Application (OSSA), support for visualization development has been modest, particularly within the Earth Science and Application Division (ESAD) which supports the EOS project. Still, development activities within the Informations Systems Branch (Code SMI) and the various science branches of the ESAD and the Solar

System Exploration Division have resulted in successful tools, capable of meeting a limited part of the EOS requirements.

Unfortunately, without proper direction and appropriate support, these efforts will be inadequate for meeting the visualization requirements within EOS. At present, visualization development activities within OSSA and the ESAD can be characterized as:

- Fragmented, with little overall direction or coordination
- Lacking an established and understandable organizational structure for funding both general tool and application-specific development
- Lacking a firm commitment or plan for meeting the true visualization needs of the EOS science community
- Lacking adequate mechanisms for technology transfer both within OSSA and between OSSA and OAST
- Experiencing inefficient use of limited funding

This environment is resulting in inefficient development which does not adequately address the full needs of the scientist. In particular, there is much redundant programming and little effort directed toward creating tools that can be easily extended and customized. Because of archaic programming practices, functionality developed within one program is not easily transferred into other development efforts. Attempts to fulfill application-specific requirements often result in large-scale application development rather than the smaller-scale addition of modules to existing applications. Finally, the absence of support for small-scale application development within science proposals generally results in the use of inadequate tools to meet science objectives, or the scavenging of science funds in order to meet development requirements.

Required in-house development is of two types: (1) development of general tools to meet the needs of a wide range of users not being met by existing software, and (2) extension, modification, or integration of existing tools to meet application-specific requirements. Unfortunately, the lack of extensible software in the past, and the lack of coordinated development activities at NASA, has not allowed a clear distinction between these efforts. Distinguishing between these efforts, and providing appropriate support for both, is vital to meeting present and future visualization requirements in a cost-effective manner.

While development of general visualization tools should be intimately related to the scientific objectives of the EOS program, the current philosophy that most visualization development should be "piggy-backed" with science proposals has been ineffective. The reasons for this are:

- Scientists are reluctant to include software development cost in science proposals for fear that the proposal costs will not be competitive;
- If only 80% of the contract is funded, software development is usually the first to be cut; the scientist will then fall back to available tools that are probably inadequate for his/her needs;
- Tying development intimately to a particular science proposal generally results in a very specialized and hard-wired tool, and does not encourage general, reusable tools for meeting other science requirements;
- Strictly creating visualization tools required for a specific science proposal does not address nor encourage the exploratory development of new techniques required to advance visualization to a state required to meet the EOS objectives.

It is recommended that general tool development be separated from specific science proposals, but be conducted by teams consisting of both computer specialists and scientists. In contrast, application-specific development could be directly tied to particular science objectives, but should receive, at least, minimal funding earmarked for that effort.

DIAMONDS AND DINOSAURS

A difficult challenge for NASA will be that of recognizing and supporting in-house "diamonds" (experimental development efforts which hold much promise for scientists) and of preventing, recognizing, and dealing with "dinosaurs" (old development programs with limited momentum and decreasing application). Diamonds typically begin as simple experimental ideas that, through proper development and strong acceptance by scientists, develop into applications with significant potential. However, without proper nurturing, such programs quickly falter due to stagnation, or more commonly, as a result of becoming overwhelmed with the responsibilities associated with distributing and supporting a successful application program. Without a proper mechanism in place to recognize these diamonds, and to assist in the transition of these development efforts, continued development and upgrading of these tools is sacrificed for the sake of maintenance and support, and the science community losses a potentially valuable tool.

At the other extreme are the large programs with a long history of development and maturity which in many cases offer significantly more functionality than other existing programs, but which have become quite archaic with regard to their interfaces, program structure, data formats, graphics capabilities, and portability to modern computing platforms. This is primarily true with regard to image processing applications. Total reprogramming of this functionality using a more efficient, extensible, and portable approach, is an ambitious effort that has traditionally been avoided. However, simple porting to newer architectures is inadequate. NASA must evaluate the importance of the functionality that exists within these programs and determine if and how it intends to incorporate this functionality into future applications.

Preventing, or at least slowing, the creation of future dinosaurs is also an important issue. In fact, the issues involving diamonds and dinosaurs are not unrelated. The main factors affecting the creation of dinosaurs from diamonds are initial design, project transitions, upgrading, and perhaps, inevitability. Certainly, programs that are less flexible to extension and modification, or are too intimately tied to data structures, have a greater danger of rapidly becoming future dinosaurs. Again, the principles of abstraction and modularity inherent in object oriented programming could greatly increase the useful life of programs which adhere to these principles. Secondly, it is important that NASA properly assist the transition of the program through its various stages, while providing adequate funding for both user support and program upgrading. Still, while these measures may significantly extend the useful life of future development efforts, it is probably inevitable that any application will, at some point in time, reach a point of questionable usefulness. NASA should also have a proper mechanism in place for phasing out continued support, while minimizing inconveniences to on-going users.

MAIN AREAS FOR CONSIDERATION

Based on these findings, the following general areas have been recognized as requiring consideration:

- **Integration of Visualization With Data Management and Analysis**
- **Application Development for the Scientist as End-User**
- **Providing Extensibility Without Complexity**
- **Application-Specific Programming Without Redundant Programming**
- **Simplification of Complex Heterogeneous Environments**
- **Development of New Techniques and Components**
- **Education and Communication**
- **Distribution and Maintenance**

These areas of concern will be discussed briefly below. More specific deficiencies and requirements will be discussed in a follow-up report. Recognition of these areas of required concentration does not, by default, imply that NASA will need to take full responsibility for meeting these needs. In fact, meeting these needs will require a proper balance of efforts and adequate communication between NASA, vendors, contractors, and other outside developers.

Integration of Visualization With Data Management and Analysis

As illustrated in Figure 2, data management, data analysis, and data visualization should be viewed as a triad of functions which can act on scientific data. In order for this triad to operate effectively, there must be adequate data paths between all components and they must be bidirectional. At present, some tentative links exist within this triad, while others are nonexistent. In particular, the visualization function presently serves primarily as a backend process, with little return interaction with analytical tools or database management.

In particular, the links between visualization and analysis are critical, but are, at present, inadequate. The application of visualization techniques can provide significant insight into scientific problems, but it is vital to the scientist to test and demonstrate this insight using

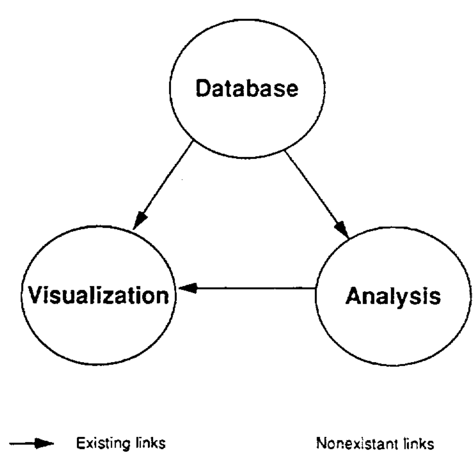


Figure 2. Triad between Data Management, Analysis, and Visualization functions, showing existing and nonexitant paths.

analytical methods. Until the scientist can easily and interactively go back and forth between the visual and analytical domains, visualization will not be considered a true scientific tool by the majority of scientists.

Although data sets derived from analytical or visualization tools may be fed back into the database, the data management system still essentially acts as a frontend entity which provides data to be used by visualization and analysis tools, but otherwise interacts very little with these functions. Particularly with the enormous database associated with EOSDIS and the increased size of many data sets, it is becoming more important for visualization and analysis tools to be able to query the data management system regarding content of the data. As an example, a visualization tool should be able to request of the data management system, "Show me browse images of other data sets for the area in this bounding box, where the ratio of this parameter to this parameter is greater than this value." Similarly, numerical models should be able to request additional data from the database based on immediate results obtained during the model runs.

These examples imply more sophisticated data management capabilities and database than those required for simple inquiries based on sensor type, date of acquisition, and nadir location. They also suggest a different level of interaction between the client (in this case, another software application) and the data management system, than might be the case if the client were an actual user.

Finally, navigation, or knowing the spatial and temporal location of each data point, is a critical issue that spans all components of this triad. The increasing requirement to combine data sets from a variety of sensors demands increased flexibility for navigation of the data within all three components. Data Management requires the ability to subset data based on precise spatial and temporal parameters, visualization requires the display of several data sets within a variety of projections, while analysis requires access to low-level data, as well as the ability to grid data into common grids. The present tendency to create higher level data sets, consisting of regularly gridded data in standard projections, only partly alleviates this challenge. Because each data set may have different scientific or instrument requirements, consensus with regard to standard grids and projections is difficult and often results in corruption and decreased scientific validity of the data. Regridding previously gridded data results in dramatically increased errors. It is therefore best to delay gridding and projection til as late in the visualization and analysis process as possible. This requires the availability of smart navigation tools, as well as the accessibility to low-level navigation information within the data sets.

Application Development for the Scientist as End-User

As discussed previously, the most immediate need for visualization is that of getting useful and usable interactive tools into the hands of the scientist. This requires that the tool meet the actual needs of the scientist, and not force the scientist to adapt to tools that are inappropriate for his/her needs. This also requires the operation of the tool, as well as data input and output of results, be intuitive and simple to the scientist. Interfaces should be simple and should speak the scientist's language where possible. The operation of the tool should be intuitive and logical to the scientist, rather than requiring the scientist to adapt to the logic of a computer graphics expert. In essence, the scientist should be able to focus on the analytical and visual results, and not on the operation of the tool.

Few, if any, of the available visualization tools meet these needs, at present. Those which are simple to use are often inflexible and don't therefore meet required scientific objectives. Those which are extensible are often too complex to operate and don't, at present, allow adequate means of simplifying this complexity.

Providing Extensibility Without Complexity

Many COTS visualization programs have not been well accepted into the EOS scientific community because of the lack of extensibility and flexibility. Without the ability for the scientist or computer specialist to input data in new formats, or to extend the functionality of the program with user-defined modules, the scientist is often forced to (1) adapt his/her data and needs to fit into visualization tools which are inappropriate for analyzing the data, generally with poor results, (2) apply several independent visualization tools to the data in order to analyze different aspects of the data, (3) give up using present visualization tools in lieu of more traditional, but inadequate data analysis tools, or (4) initiate an in-house development effort which is costly in time and funding, with probably 50 to 75% of the effort applied to redundant programming of available capabilities.

One of the main reasons quoted by scientists for initiating an in-house development effort has been the need for extensibility and flexibility. This has also been a major reason for the success of many NASA-developed and other public-domain application programs, since source code is generally provided for these programs. Furthermore, the COTS visualization programs that have gained the most acceptance within the EOS science community are those which provide a high degree of extensibility and flexibility, including AVS, SGI's Explorer, NCAR graphics, PV-Wave, and IDL. The development of these tools, as well as public domain tools such as Khoros, are major milestones in the advancement of scientific visualization.

However, the generality and extensibility of these tools have been generally accompanied by a high degree of complexity of use. PV-Wave, IDL, and NCAR graphics require custom programming and command-line interfaces, while the graphical programming paradigms of AVS and SGI Explorer have given rise to the self-describing term, "button hell". These complexities, at present, limit the use of these programs to those scientists who have the time and initiative to become highly involved in the design of their tools. In order for these tools to be applicable by a high percentage of EOS scientists, it is necessary that computer specialists be able to adapt these tools to meet the particular needs of a scientist or project while hiding the complexities of the extensible environment from the scientist using a simple and intuitive user interface.

PV-Wave and IDL have made great strides in allowing this, although these programs are primarily limited to interacting in 2D space. The developers of AVS and SGI Explorer are likewise aware of the issue, but must work out concerns regarding adaptive scripting, decision nodes, and user-defined GUIs before these programs will meet the needs of a large percentage of EOS scientists. Of course, any in-house or other NASA-funded development should also concerned with extensibility while allowing for simplification and customization.

Application-Specific Programming Without Redundant Programming

Redundant in-house programming within the Earth systems science community is extensive. The primary reason given for abandoning OTS software in favor of in-house develop is because available OTS software (a) did not do what was required for specific applications and lacked extensibility and adaptability to modify it accordingly, (b) had a user interface that was inappropriate or too cumbersome, (c) was too expensive, or (d) was inefficient in its use of computer resources. Although most of these in-house application programs provide important features not available in other packages, each of these development efforts probably contain 50-75% redundantly programmed code. In other words, in order to provide the new application-specific functionality, it was necessary to also redundantly develop functionality that was already available in several other programs. Cannibalism of previously developed code has been helpful in some circumstances, but most of this code is too intimately tied to the environment of the original program, or limits the flexibility or functionality of the new program. Considering that most of these in-house development teams consist of 1-3 programmers, a considerable amount of effort is concentrated in programming functionality that already exist.

As discussed above, the availability of more extensible and flexible visualization tools in the future may minimize, to some extent, the amount of in-house programming, and thus the amount of redundant programming. The extent to which these and other OTS programming environment will meet our future development needs is still uncertain. However, there will probably still be requirements in the future for in-house application-specific programming. The remaining problem then becomes how to reduce the amount of redundant programming, while meeting the requirement for application-specific programming within the EOS community. The answer probably lies in a combination of hierarchical and object-oriented development.

The benefits of object-oriented programming include the reusability of modular components, and the minimization, through abstraction, of complexities resulting from our heterogeneous computing environment. Hierarchical development implies that general components, which provide functionality common in many programs, be developed and made available to application-specific development efforts, thereby allowing these efforts to concentrate on providing value-added programming rather than redundant programming.

Simplification of Complex Heterogeneous Environments

Scientists and computer specialist are continuously required to operate in a complex heterogeneous computing environment. Data and applications exist on several computers running DOS, OS/2, Mac, Unix, VMS, or MVS operating systems, as well as having various graphics protocols and functionality. Data files reside on an array of storage devices in a wide variety of standard and proprietary data formats. Networks operate under a mix of protocols, with each protocol or hardware platform having its own peculiarities with regard to file transfer. Access to hardcopy or video devices is through serial, parallel, ethernet, or SCSI ports, with each port and each device requiring different protocols or file formats. The need to operate within this complex environment accounts for a significant loss of efficiency in the use of our computing tools. Several scientists have estimated that, at least, 60-70% of their computing efforts are consumed in dealing with the complexities of networks, data formats, and operating systems, leaving very little time and energy for concentrating on scientific analysis.

There are two potential solutions: homogenize the computing environment or make working within the heterogeneous environment simple and transparent. The latter is more realistic for both the present and the future. While it is important to develop and adhere to common data formats, as well as standards in operating systems, network protocols, windowing environments, and graphics languages, it unreasonable to believe that our computing environment could ever become, and remain, homogeneous. Therefore, in order to make efficient use of our computing resources, it is vital that operating within this complex environment become simple and less painful.

In essence, a scientist should be able to pull an icon representing data over to an application icon, and have the program operate regardless of underlying complexities. The scientist sitting on a Unix graphics workstation, for example, should not be required to deal with underlying complexities resulting from the possibility that the data resides on a different computer, in a format incompatible with the application program, which possibly resides on a VMS-based computer. Furthermore, the scientist should be able to request a color hardcopy or video animation, with little concern as to the underlying complexities involved in accomplishing that request.

There is no technical reason why such a solution cannot exist today. In fact, present efforts to provide Unix and DOS with a more Mac-like interface, as well as efforts directed toward object oriented operating systems, are all major advances toward this goal. Still much emphasis and effort is required, in order to move the scientific computing community into such an environment.

Development of New Techniques and Components

Although many of our immediate needs involve integrating present capabilities into useful and usable application for scientists, there is still a need to continue investigation and development of new techniques and components. In a few cases, this may involve a new approach to analyzing and visualizing data and require development of a complete application program. The NASA funded programs, Linkwinds, Interactive Image Spreadsheet (IISS), and VisAD, are examples of such.

However, in most cases, the development of new techniques should not require development of a complete application in order to be applied. Instead, newer techniques, such as spectral mixing, volume rendering algorithms, or accurate navigation routines, for example, should instead be developed as functional components or modules which can be added to existing applications. Unfortunately, the lack of module programming and coordination of development in the past, has generally resulted in the development a new application around each new technique. Adherence to object-oriented programming principles and the use of extensible software would maximize our future development efforts and minimize the cost of leading edge development.

Education and Communication

Many of the present challenges in placing appropriate visualization tools into the scientist's hands are the result of a simple lack of communication and education. Many scientists lack an understanding and appreciation of the importance of visualization in meeting their scientific objectives, while others lack an understanding of the present limitations. Project managers aren't certain whether visualization needs exists and if these needs should be met with COTS software or in-house development. Many developers lack a true understanding of the needs of the scientists and are often frustrated when scientists don't appreciate the tools presented to them. In particular, hardware and software vendors must rely on their own restricted capabilities for determining the present and future needs of EOS scientists, generally obtaining a narrow view based on contact with a few select individuals. Even within NASA, itself, scientists, developers, and project managers are often unaware that efforts going on at other centers or within other organizations, might be beneficial and applicable to their own activities.

Furthermore, scientists and specialists, alike, are often unaware of the existence and functionality of many of the available tools that might be useful to them. It is a very time-consuming process to search out a specific tool that might serve one's need, to download or order it, install it on the hardware platform at hand, learn to operate it, get appropriate data into it, and then to test it for applicability to the scientific problem, at hand. If that tool does not meet the total needs, or in some cases does not meet the needs at all, then the search and testing process must begin again or be abandoned. Scientists and visualization specialists would benefit greatly from a system that not only informed them of the availability of appropriate visualization tools, but one that also provided them unbiased opinion and expert technical assistance in evaluating whether a given tool would meet their objectives. More than simply a repository of programs and data, scientists need knowledgeable assistance.

Communication must improve both within NASA and between NASA and its contractors and vendors. Often, a lack of communication is responsible for NASA unnecessarily "contracting out for expertise", rather than using expertise that already exist within other organizations of NASA. EOS scientists must be made aware of tools that are available to them, while developers must become better informed of the immediate and future needs of the scientists. Particularly if COTS software is to play a major role in meeting future needs, then vendors must become more of a true partner in the process.

Distribution and Maintenance

The distribution and maintenance of application software is a significant challenge, particularly with regard to NASA funded or public domain software. As an in-house development project begins to reach a certain level of success, it is often faced with the dilemma of limiting the distribution and support of the software, or redirecting funding and manpower from development to support. Without the proper funding and mechanism for assisting in the transition from experimental to operational, distribution and user support can become a serious drain on the project effort. This results in either poor user support or in stagnation of development and growth of the program. Under such conditions, younger programs often never reach a useful level of maturity, while older efforts become archaic do to lack of program modification and upgrading.

In essence, there is a need for assistance in transitioning experimental software into operational phases, primarily for distribution and user-support efforts. It is important that distribution and support efforts be established well before a program becomes completely polished and documented. In addition, there should be a mechanism for distributing unsupported or modular code, since these are often of significant use to in-house developers.

In addition, there are many public domain applications which can be of significant benefit to the EOS scientific community. If these efforts do not continue to receive adequate funding for product distribution, user support, and continued development, these projects are forced to either dissolve or go commercial, often with disastrous results for the user and development alike. For applications found useful for meeting NASA's scientific objectives, it is generally more cost-effective to provide these projects some level of support through licensing or funding, than to replace the functionality of these program with in-house development.

For COTS software, the major challenges lie primarily in the ability to test these programs before buying, and in establishing proper mechanisms for licensing useful software within a distributed environment. Improved site licensing mechanisms, such as network licenses, floating licenses and "pay-as-you-use" licenses, would be useful. In addition, rapid procurement of useful, tested COTS software is essential.

CONCLUSIONS

Visualization is vital for meeting the scientific objectives of the EOS mission. Although our present suite of visualization techniques is impressive and powerful, the application software for putting these capabilities into the hands of the scientist is, at present, inadequate. The development of extensible, user-friendly, object-oriented software within the commercial software industry, is helping to increase the probability that COTS software can serve as a core for meeting many visualization needs. However, there will still be a need for in-house efforts directed at extending and modifying these tools to meet application-specific requirements, and for developing leading edge techniques and applications not available from the commercial front.

A follow-up report will investigate viable options for meeting the visualization requirements of the EOS mission in the future. These options will focus on two major objectives: increasing the effective use of visualization tools by scientists, and maximizing the return on development efforts.

INFORMATION SYSTEMS RESEARCH AND TECHNOLOGY: PROGRAM REVIEW

Mr. Glenn Mucklow
NASA Headquarters, OSSA

August 13, 1992

INFORMATION SYSTEMS BRANCH — SSA

Information Systems Research & Technology

Information Systems Research and Technology

Program Overview

Office of Space Science and Applications,
Flight Systems Division
July 29, 1992 -- Chart 1

INFORMATION SYSTEMS BRANCH — SSA

Information Systems Research & Technology
The Two Program Objectives

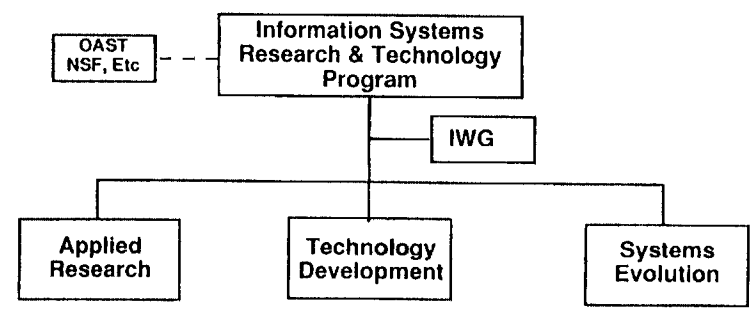
Apply advanced information systems technology, as appropriate, to improve support to OSSA science programs

Enable continual evolution of OSSA data systems environment and supporting infrastructure

Office of Space Science and Applications,
Flight Systems Division
July 29, 1992 -- Chart 2

INFORMATION SYSTEMS BRANCH — SSA

Information Systems Research & Technology
Program Elements



Office of Space Science and Applications,
Flight Systems Division
July 29, 1992 -- Chart 3

Information Systems Research & Technology
Applied Research Element and Approach

The Applied Research element

applies

appropriate computer and information systems science and technology from universities, government, and industry

to the OSSA Earth and space science process

through peer-reviewed research in

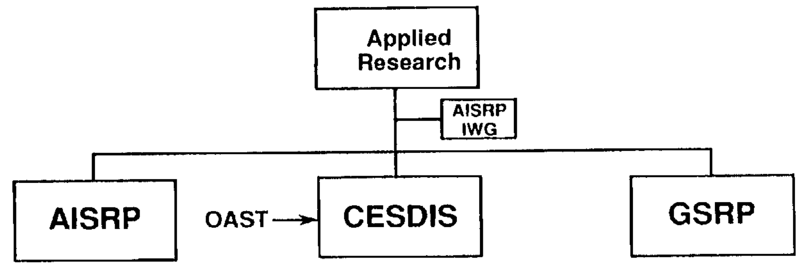
the Applied Information Systems Research Program (AISRP),

the Center of Excellence in Space Data and Information Sciences (CESDIS),

and participation in the Graduate Student Research Program (GSRP).

Office of Space Science and Applications, Flight Systems Division
July 29, 1992 -- Chart 4

Information Systems Research & Technology
Applied Research Element



Office of Space Science and Applications, Flight Systems Division
July 29, 1992 -- Chart 5

Information Systems Research & Technology
Technology Development Element and Approach

The Technology Development element

enables

continual evolution

of OSSA data systems environment and supporting infrastructure

through development and demonstration of

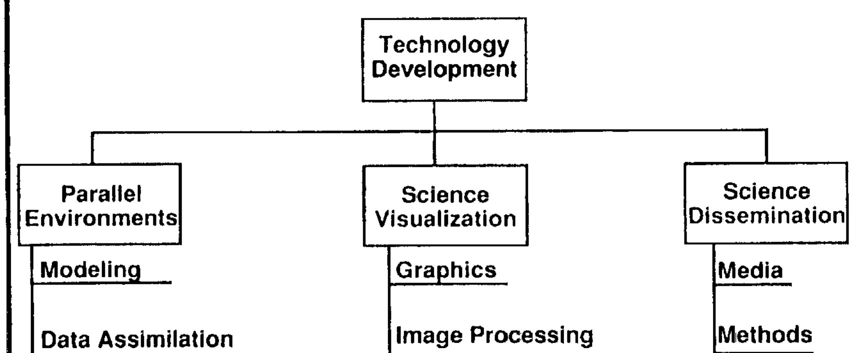
Science Visualization tools for image processing and graphical techniques,

Parallel SW Environments for model and data assimilation, and

Dissemination techniques for faster and better publication of results.

Office of Space Science and Applications, Flight Systems Division
July 29, 1992 -- Chart 6

Information Systems Research & Technology
Technology Development Element



Office of Space Science and Applications, Flight Systems Division
July 29, 1992 -- Chart 7

INFORMATION SYSTEMS BRANCH

Information Systems Research & Technology
Systems Evolution Element and Approach

The Systems Evolution element

facilitates

testing, evaluating, inserting, and maintaining current information systems technology

in the OSSA Earth and space science program environment

through development of OSSA-wide

Policies,

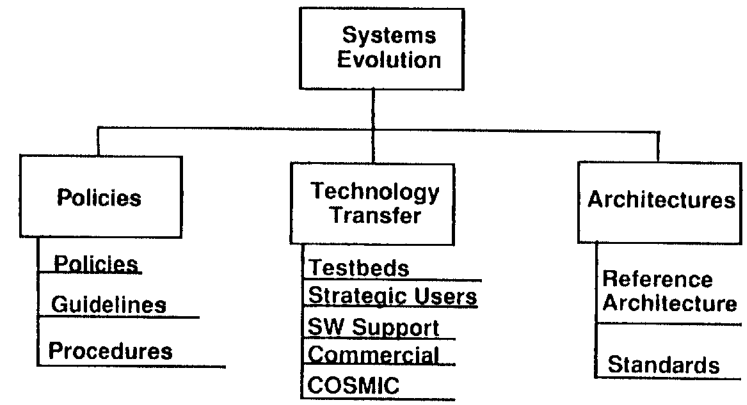
Architectural Guidelines, and

Technology Transfer Infrastructure.

Office of Space Science and Applications, Flight Systems Division
July 29, 1992 -- Chart 8

INFORMATION SYSTEMS BRANCH

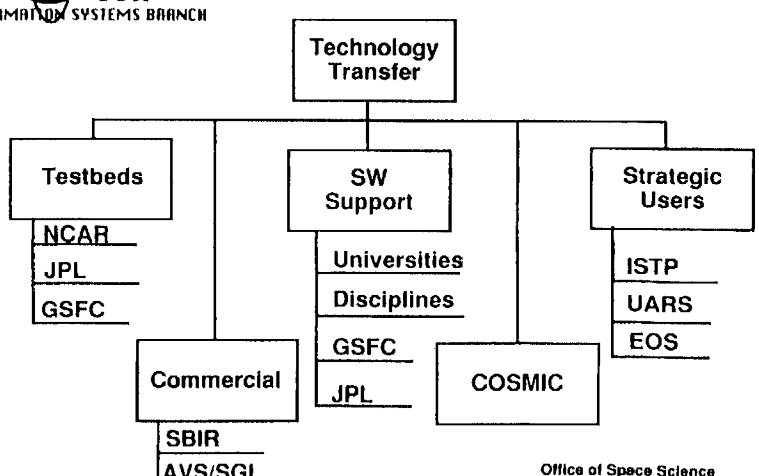
Information Systems Research & Technology
Systems Evolution Element



Office of Space Science and Applications, Flight Systems Division
July 29, 1992 -- Chart 9

INFORMATION SYSTEMS BRANCH

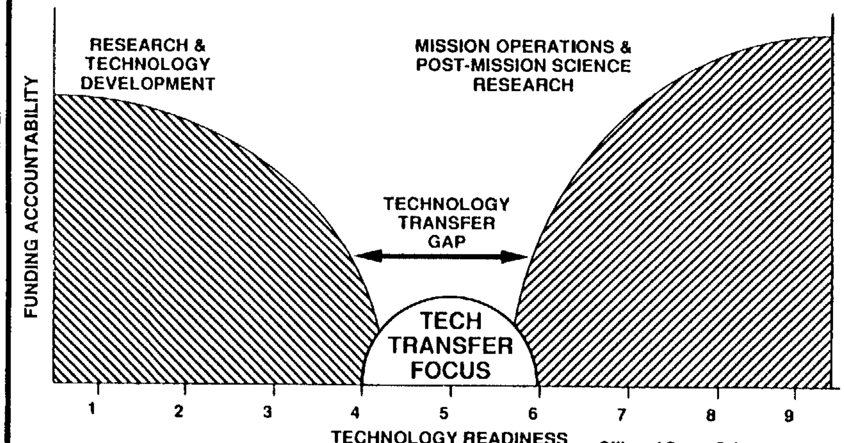
Information Systems Research & Technology
Technology Transfer Program



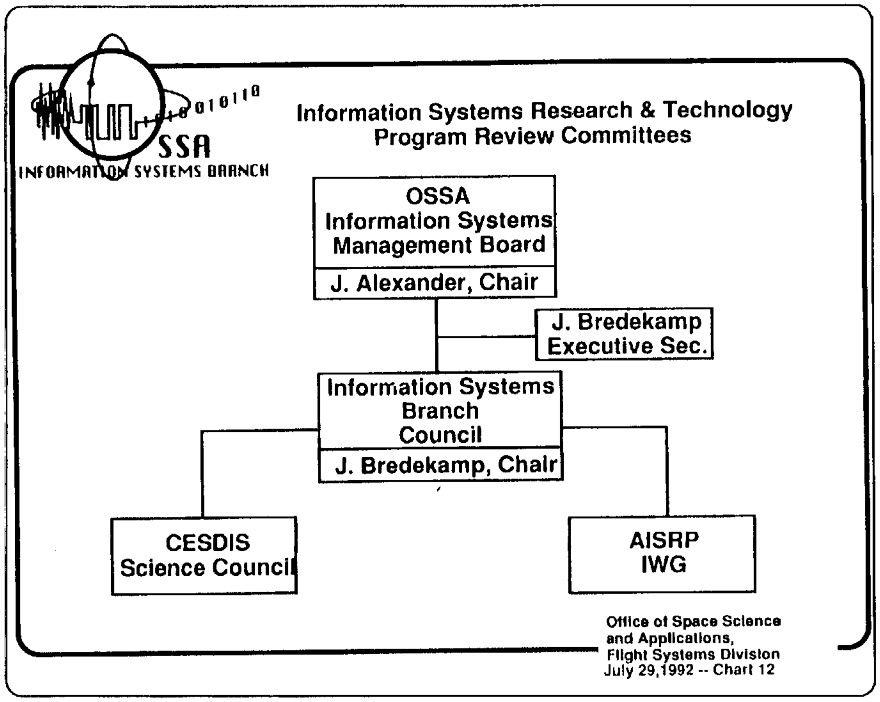
Office of Space Science and Applications, Flight Systems Division
July 29, 1992 -- Chart 10

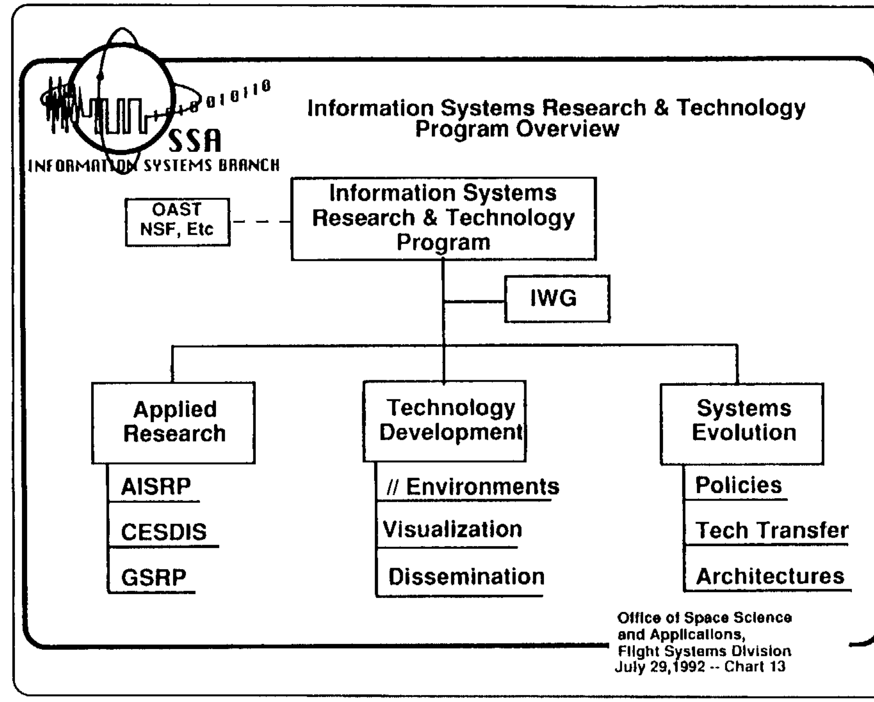
INFORMATION SYSTEMS BRANCH

Information Systems Research & Technology
Technology Transfer Approach-- Bridge the Gap



Office of Space Science and Applications, Flight Systems Division
July 29, 1992 -- Chart 11


INFORMATION SYSTEMS BRANCH
SSA
Information Systems Research & Technology
Program Review Committees
OSSA
Information Systems
Management Board
J. Alexander, Chair
J. Bredekamp
Executive Sec.
Information Systems
Branch
Council
J. Bredekamp, Chair
CESDIS
Science Council
AISRP
IWG
Office of Space Science
and Applications,
Flight Systems Division
July 29,1992 -- Chart 12


INFORMATION SYSTEMS BRANCH
SSA
Information Systems Research & Technology
Program Overview
OAST
NSF, Etc
Information Systems
Research & Technology
Program
IWG
Applied
Research
AISRP
CESDIS
GSRP
Technology
Development
// Environments
Visualization
Dissemination
Systems
Evolution
Policies
Tech Transfer
Architectures
Office of Space Science
and Applications,
Flight Systems Division
July 29,1992 -- Chart 13

ELECTRONIC MAIL AND NEWS GROUPS

Mr. Wrandle Barth
Goddard Space Flight Center

August 13, 1992

Using the NASA POBox

HUGHES
HUGHES STX CORPORATION

- Uniform, deducible names (usually first initial and last name) regardless of recipient's system
- Syntax depends only on sender's system:
  - Internet . . . . . . aisrp-gmucklow@pobox.nasa.gov
  - NSI-DECnet . . . EAST::"aisrp-gmucklow@pobox.nasa.gov"
  - BITNET . . . . . . aisrp-gmucklow@pobox.nasa.gov
  - NASAMAIL . . . (SITE:INTERNET,ID:<aisrp-gmucklow(a)pobox.nasa.gov>)
  - JSC PROFS . . . TO: NASAMAIL(POSTMAN)
    first line of text:
    TO: aisrp-gmucklow@pobox.nasa.gov
- Distribution lists for groups:
  - aisrp-members . . . for all members
  - aisrp-pi . . . . . . for all PIs
  - aisrp-general . . . for general discussions

On-going Electronic Discussion Groups Via Central Bulletin Board

HUGHES
HUGHES STX CORPORATION

- Minimizes storage
- One way to provide central library of information
- Usually command-line interface; specific to system
- Mainframe response variability
- Can "disappoint" if discussion frequency varies

Discussion Groups via USENet News

HUGHES
HUGHES STX CORPORATION

- Good for wide-open, self-subscribed discussions
- If not open to all, must arrange for feed from other contact
- A variety of news readers (can ftp from dftsrv.gsfc.nasa.gov)
  - UNIX . . rn, nn, tin, trn, gnews, vn
  - VMS . . . vnews
  - PC . . . . trumpet
  - Mac . . . TheNews, Newswatcher, TCP/Connect, nuntius and various hypercard readers
- No extra effort if already reading news
- Groups categorized within:
  - comp . . computer science
  - sci . . . . technical discussions, other science
  - misc . . . hard to categorize, or multiple categories
  - soc . . . . social issues, socializing
  - talk . . . debate, controversy
  - news . . discussions about USENET News itself
  - rec . . . . recreation and hobbies

Discussion Groups via Electronic Mail

HUGHES
HUGHES STX CORPORATION

- Discussions delivered into mail system
- Using LISTSERV, users can subscribe, unsubscribe any group
- Additional discussion groups can be created for special needs
- Use familiar commands of your mail system to reply, post, print, file, etc.
- Works even if behind mail gateway
- To subscribe/unsubscribe: mail to `listserv@pobox.nasa.gov` with body

  `subscribe   aisrp-general` *first-name last-name*
  `unsubscribe aisrp-general`

- To post: mail to `aisrp-general@pobox.nasa.gov`

Contact registrar to have your information added or modified in the AISRP directory:

`aisrp-registrar@pobox.nasa.gov`

NOTE: see back panel for proper syntax from your system

To join in the AISRP general discussion group, send a mail item to

`listserv@pobox.nasa.gov`

and as the only line of the body, type:

`subscribe aisrp-general` *fn ln*

where *fn* is your first name and *ln* is your last name.

Contact the NASA Science Internet Help Desk for questions on the use of the NASA POBOX:

NSI Help Desk
(301) 286-7251
helpdesk@pobox.nasa.gov

POBox Syntax

To send mail to a user or group, chose the syntax below based on the system you are sending from. Note that "aisrp-" is optional when the name is unique within the NASA POBox system. *nnn* is the user's first initial and last name, or one of the groups "members", "pi", or "general".

From Internet

aisrp-*nnn*@pobox.nasa.gov

From NSI DECnet (formerly SPAN)

EAST::"aisrp-*nnn*@pobox.nasa.gov"

From BITNET

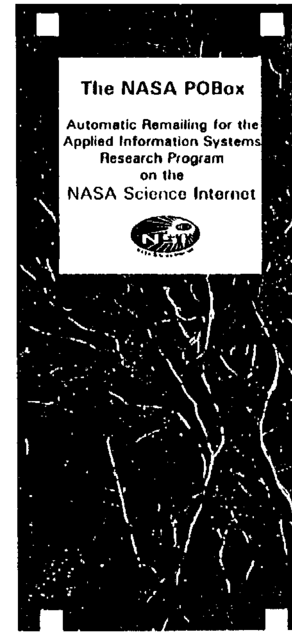
aisrp-*nnn*@pobox.nasa.gov
(for most sites; may be stored as a NICKNAME)

From NASAMAIL

(SITE:INTERNET,
ID:<aisrp-*nnn*(a)pobox.nasa.gov>)

From JSC PROFS

TO: NASAMAIL(POSTMAN)
First line of text:
TO: aisrp-*nnn*@pobox.nasa.gov



The NASA POBox

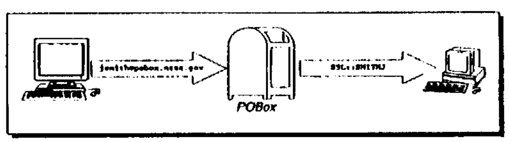
Simplifies mail exchange among AISRP researchers on various systems:

- The Internet
  including NSI (the NASA Science Internet), NSFNet, US regionals, and international components
- NSI DECnet (formerly SPAN)
  with interconnections to HEPNET, THEnet, GEONET
- RSCS Networks
  including BITNET, CDNN, EARN, NETNORTH
- X.25 Networks/Nodes
  including ACCUNET, OSFCMAIL, NASAMAIL, NPSS, OMNET, SPRINTMAIL, TYMNET
- LAN & personal computer systems connected to the above

All capabilities directly integrated into your existing mail system:

- No separate systems to log into
- No new editors to learn
- All the printing, searching, and filing capabilities you use for other correspondence

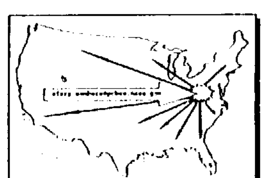
Mail Forwarding to an Alias:



- Uniform, deducible names regardless of recipient's system
- Remains constant when recipient's end system changes
- Passes through gateway systems transparently
- Supports your mail system's cc:, reply, and forwarding capability
- Natural evolution to X.500

Mail Forwarding with Distribution List Explosion

- Used to support distribution of AISRP electronic newsletters and bulletins
- Electronic discussion groups supported
- Mail sent to list name is forwarded to all those on list
- Three lists initially:
  `aisrp-members`: all members
  `aisrp-pi`: all PIs
  `aisrp-general`: general discussion group for those who wish to subscribe



NASA SCIENCE INTERNET DEVELOPMENTS

Dr. Christine M. Falsetti
Ames Research Center

August 13, 1992

NASA Science Internet Developments
presented to the
Applied Information Systems Research Program Workshop II

August 1992

Christine M. Falsetti
Project Manager
NASA Science Internet Office
Information and Communications Systems Division
AMES RESEARCH CENTER

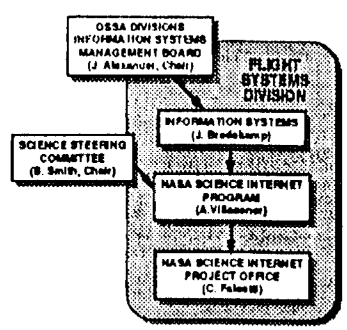
NASA

The NASA Science Internet

Recognizing that science communications networking is an integral element of successful science, the NASA Science Internet Office was established in 1988 to provide communications to NASA's Office of Space Science and Applications, OSSA.

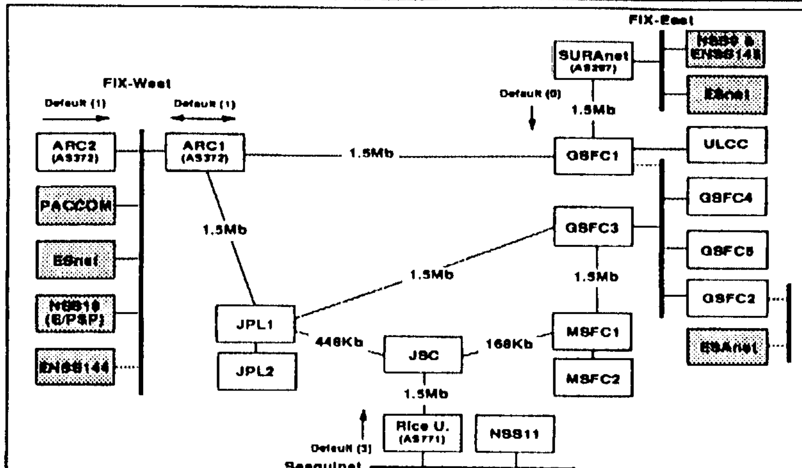
The NSI provides computer networking services, management and operations support, and technical assistance to authorized users throughout NASA centers and research institutions worldwide.

NSI 's goal is to provide a high-speed communications network that connects all space scientists, providing ready access to data and information stored anywhere in the world.



NSIC Falsetti 8/4/92 2

NSI Backbone



NSIC Falsetti 8/4/92 3

Current Telecommunications Infrastructure

DEDICATED POINT-TO-POINT LEASED LINES

- Local loop provided by Local Exchange Carriers (LEC's)
- Long haul provided by Inter-eXchange Carriers (IXC's)
- No customer switching equipment at TelCo central offices
- Possibility of link failure requires redundancy
- Dedicated to single user use (and billing!)
- Limited bandwidths available (9.6, 56, 1544 kbps, etc...)
- User must provide network monitoring and diagnostics (e.g., PSCN COMM)
- Routers at user sites perform packet switching
  - LAN interfaces (e.g., Ethernet, FDDI, etc.)
  - WAN interfaces (e.g., serial sync lines using V.35, etc.)
  - Switch packets from serial line to serial line or LAN
  - Provide network layer routing

Evolving Network-Based Applications...

MULTI-MEDIA COLLABORATIONS

- Videoconferencing from the PIs' workstations
- Embedded voice, video, movies in publications
- Distributed conferencing, worldwide

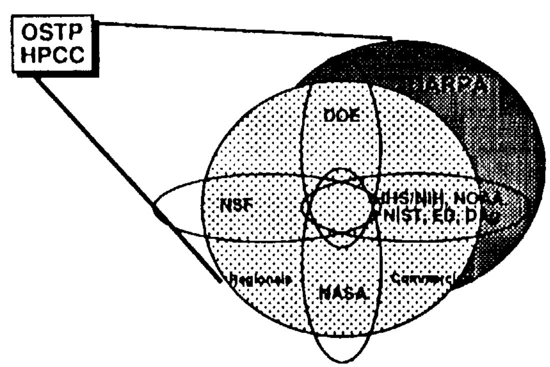
ACCESS TO DISTRIBUTED INFORMATION

- New network tools: WAIS, Gopher, Fetch, WWW, etc.

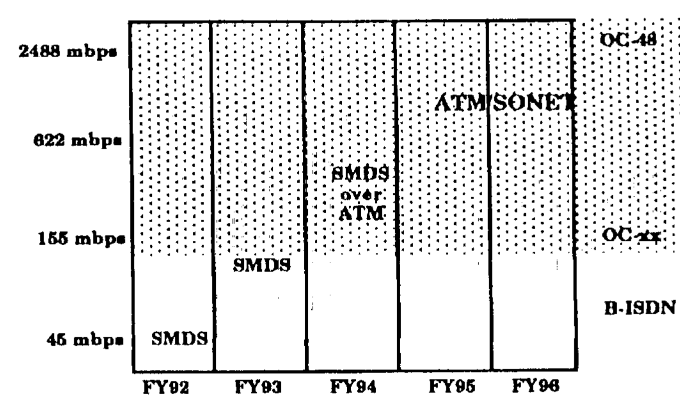
NEW TECHNOLOGY BENEFITS

- High speeds for science data (over gigabit trunks)
- Guaranteed bandwidth, on demand
- Economies of scale
- Ubiquitous access via ISDN and B-ISDN

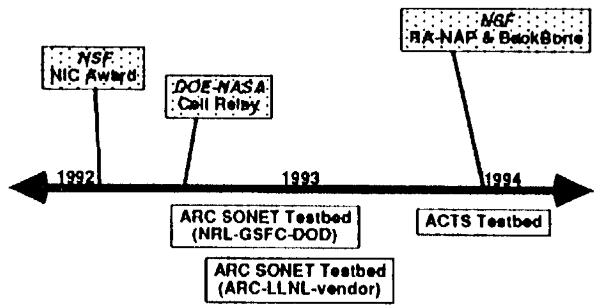
NREN Vision



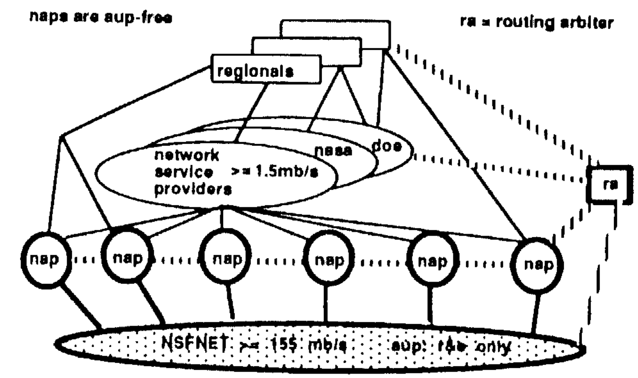
Technology Migration ==> ATM/SONET !!



NEAR-TERM NREN 1992 MILESTONES



Proposed NSFNET/IINREN Approach



NASA NREN OVERVIEW

- Focus on meeting NASA HPCC Grand Challenge requirements
- Maximizes use of existing network resources, NASA & non-NASA
- Strongly leverages on telecommunications carrier plans and investments: no private dedicated systems
- Minimizes NASA costs while still providing high performance capabilities
- Fully consistent with national program objective to construct a high performance national network infrastructure
- Provides for strong technology transfer to other programs
- Provides testbed for advanced routing and management designs for large scale public data networks

NASA-HPCC: COMPUTATIONAL AEROSCIENCES

GOAL:

Develop necessary computational technology for the numerical simulation of complete aerospace vehicles for both design optimization and analysis throughout the flight envelope.

OBJECTIVES:

- Develop multidisciplinary computational models and methods for scalable, parallel computing systems.
- Accelerate the development of computing system hardware and software technologies capable of sustaining a teraflops performance level on computational aeroscience applications.
- Demonstrate and evaluate computational methods and computer system technologies for selected aerospace vehicle and propulsion systems models on scalable, parallel computing systems.
- Transfer computational methods and computer systems technologies to aerospace and computer industries.

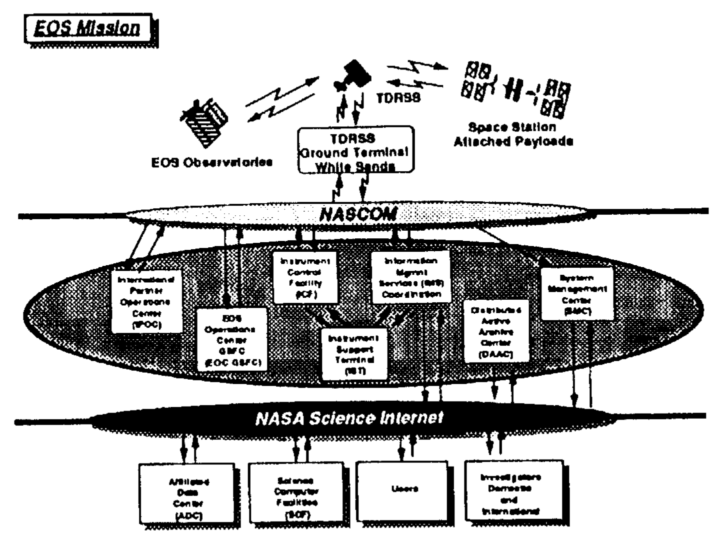
NASA-HPCC: EARTH AND SPACE SCIENCES

GOAL:

Demonstrate the potential to address the Grand Challenges afforded by teraflops systems performance on selected multidisciplinary modeling and massive data handling applications.

OBJECTIVES:

- Support the development of massively parallel scalable multidisciplinary models and data processing algorithms.
- Make available prototype scalable parallel architectures and massive data storage systems to ESS researchers.
- Prepare the software environments to facilitate scientific exploration and the sharing of information and tools.
- Develop data management tools for high-speed access, management and visualization of data with teraflops computers.
- Demonstrations of scientific and computational impact for earth and space science applications.



CESDIS: CENTER OF EXCELLENCE IN SPACE DATA AND INFORMATION SCIENCES

Dr. Ray Miller
NASA/Goddard Space Flight Center

August 13, 1992

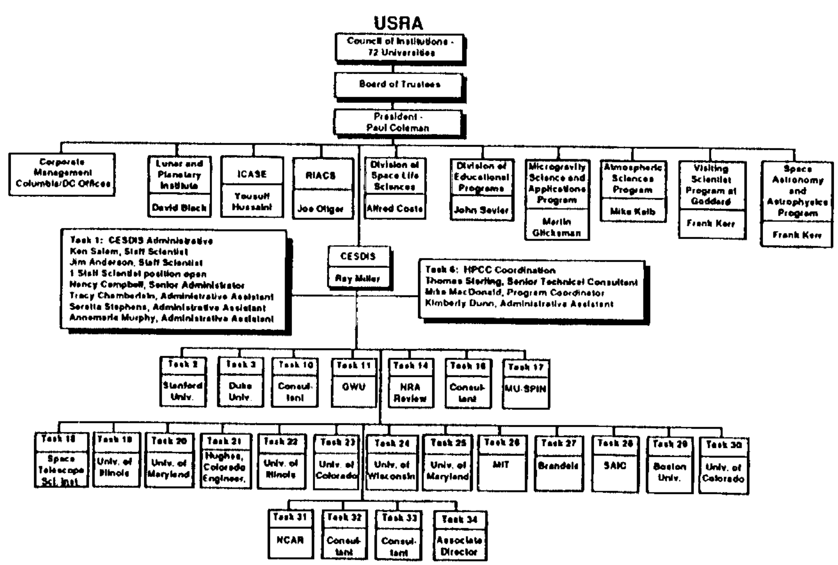
CESDIS

Center of Excellence in Space Data and Information Sciences

NASA Goddard Space Flight Center

Greenbelt, MD 20771

Operated by the Universities Space Research Association

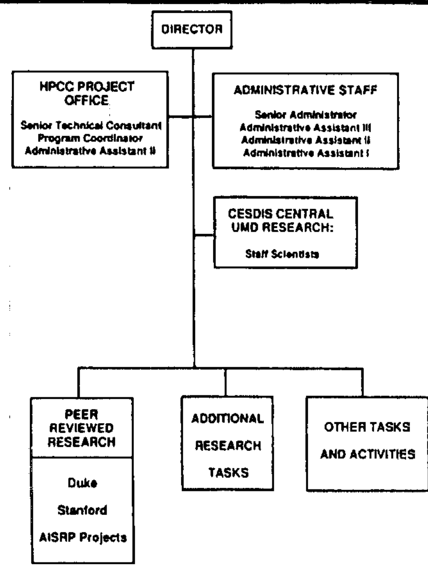


CESDIS MISSION

To bring together computer scientists from university, industrial, and government laboratories to:

- Conduct computer science research having application to Earth and space science;
- Focus attention on accessing, processing, and analyzing data from space observing systems; and
- Collaborate with NASA space and Earth scientists.

CESDIS ORGANIZATION



CESDIS CENTRAL/ UMD RESEARCH

Director: Raymond E. Miller

Graduate Assistant: Sanjoy Paul

Staff Scientist: Ken Salem

Graduate Assistants: Ahmet Sayar

Sedat Akyurek

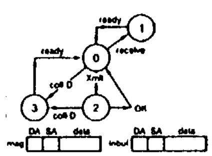
Staff Scientist: Jim Anderson

Graduate Assistant: Jae Heon Yang

COMMUNICATION PROTOCOL RESEARCH

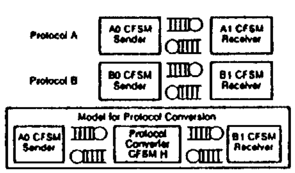
Raymond Miller - University of Maryland and CESDIS Director

Research Assistants: Sanjoy Paul, University of Maryland
Murali Rajagopal, Georgia Institute of Technology

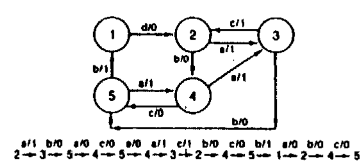


Transition	Predicate	Action
Xmit	msg ≠ ∅ ∧ medium = ∅	medium := msg; Signal(i) := transceive
OK	Signal(i) = clear	msg := ∅
coll D	medium = undefined	Signal(i) := collision
ready	Signal(i) = clear	
receive	medium DA = i	inbuf := medium; Signal(i) := transceive

SPECIFICATION AND VERIFICATION



CONVERSION



TESTING

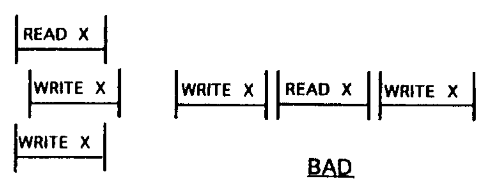
A578.001

HIGHLY CONCURRENT SYNCHRONIZATION MECHANISMS

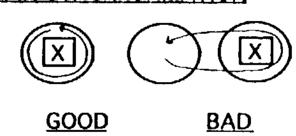
Jim Anderson - University of Maryland and CESDIS
Staff Scientist

Research Assistant: Jae-Heon Yang

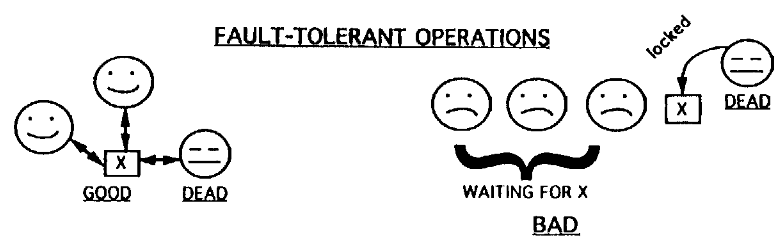
HIGHLY PARALLEL OPERATIONS



SPIN SYNCHRONIZATION

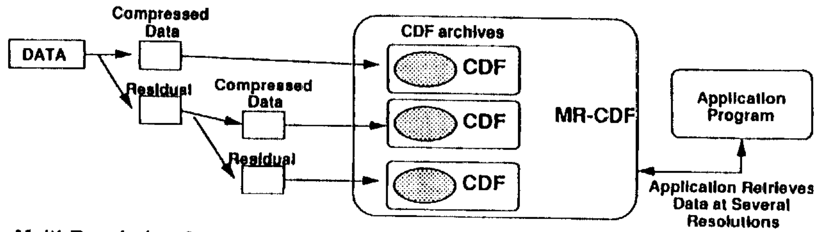


FAULT-TOLERANT OPERATIONS



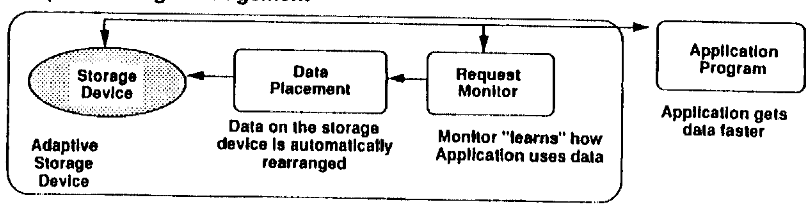
Kenneth Salem

Research Topics



Multi-Resolution Common Data Format (MR-CDF)

Adaptive Storage Management



TASK GOALS AND ACCOMPLISHMENTS

Stanford University

Computer Assisted Analysis of Auroral Images Obtained From High Altitude Polar Satellites

PERIOD OF PERFORMANCE: January 1989 - December 1992

GOALS

Develop, implement, evaluate, and utilize advanced computer software tools whose purpose is to automate the analysis of global auroral images obtained from DE-1 and Viking satellites.

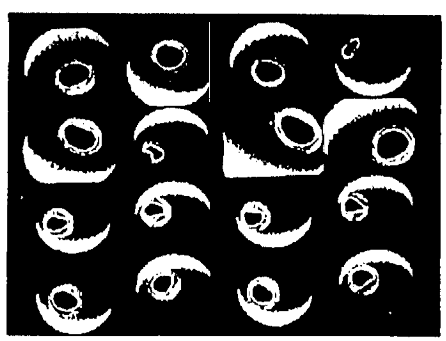
ACCOMPLISHMENTS

- Completed implementation and testing of "snakes" algorithm for finding auroral oval inner boundary.
- Developed new techniques to handle regions where daylight overlaps aurora and to simultaneously find both inner and outer boundaries.
- Demonstrated automated and interactive prototype systems to NASA scientists and installed systems on NSSDC computers.

COMPUTER ASSISTED ANALYSIS OF AURORAL IMAGES OBTAINED FROM HIGH ALTITUDE POLAR SATELLITES

CESDIS

Gio Wiederhold, C. Robert Clauer, Ramin Samadani,
Domingo Mihovilovic (Research Assistant) - Stanford University



Automatically generated inner auroral oval boundaries for 16 DE-1 satellite images applying computer-generated elastic curves or "snakes" technique.

A575 06 80110

TASK GOALS AND ACCOMPLISHMENTS

Duke University

Parallel Compression of Space and Earth Data

PERIOD OF PERFORMANCE: October 1988 - September 1993

GOALS

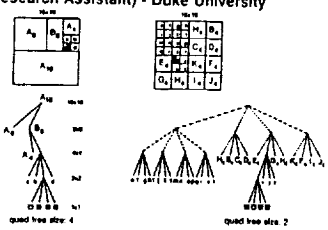
- Develop parallel algorithms and architectures for lossy and lossless data compression.
- Design compressions applicable to large variety of data.
- Design good parallel vector quantization algorithm.

ACCOMPLISHMENTS

- Devised several algorithms, investigated their theoretical aspects, and evaluated performance against USC image database to compare results with existing schemes using same images.
- Developed multiresolution lossy methods capable of controlling amount of information lost by trading off between compression rates and distortion.

PARALLEL COMPRESSION OF SPACE AND EARTH DATA

John Reif, Hillel Gazit, Tassos Markas (Research Assistant) - Duke University



The quad-tree representation of the Multi-Resolution Vector Quantization algorithm which compresses data by dividing an image block into a number of variable-size subblocks, and encoding them separately using a quad-tree representation.

AISRP PROJECTS

PROJECT	STARTING DATE
SPACE TELESCOPE SCIENCE INSTITUTE	July 1, 1992
UNIVERSITY OF ILLINOIS	June 4, 1992
TEXAS A&M	June 4, 1992
UNIVERSITY OF MARYLAND	June 1, 1992
HUGHES AIRCRAFT COMPANY	June 27, 1992
UNIVERSITY OF ILLINOIS	July 20, 1992
UNIVERSITY OF COLORADO, BOULDER	August 1, 1992
UNIVERSITY OF WISCONSIN, MADISON	July 1, 1992
UNIVERSITY OF MARYLAND	June 1, 1992
MASSACHUSETTS INSTITUTE OF TECHNOLOGY	June 1, 1992
BRANDEIS UNIVERSITY	June 1, 1992
SPACE APPLICATIONS INTERNATIONAL CORPORATION (SAIC)	July 1, 1992
BOSTON UNIVERSITY	July 1, 1992
UNIVERSITY OF COLORADO, BOULDER	June 1, 1992
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH	July 9, 1992

HIGH PERFORMANCE COMPUTING AND COMMUNICATION

PROJECT OFFICE

- To provide NASA Headquarters HPCC Project Office advice and support.
- Project Coordination
- Technical Advice/ Consultation
- Administrative Assistance
- Library Organization

Personnel:

Dr. Thomas Sterling
Mr. Michael MacDonald
Ms. Kimberly Dunn

ADDITIONAL RESEARCH TASKS

George Washington University: 9/90 - 8/91	John Sibert and Cindy Starr *Computer Graphics for Scientific Visualization* Dan Spicer
George Washington University: 10/90 - 9/92	Rainald Lohner *Computer Codes for Simulation of 3-D Compressible Magneto-Hydrodynamic Flows* Dan Spicer
George Washington University: 11/90 - 9/91	Burt Edelson and Hermann Helgert *Supercomputer Networking for Space Science Applications* ACT Satellite
Penn State University: 5/90 - 5/92	Eric Feigelson Research Assistant: Michael LaValley *Improved Data Analysis in Astrophysics*
Stanford University: 3/91 - 3/92	Philip Scherrer and Richard Bogart *AstroMail Development: an Electronic Mail System for the Astrophysics Community* Pat Gary
Brown University:	Jeffrey Vitter and Paul Howard Data Compression Algorithms

CONSULTANTS

- Noah Friedland, University of Maryland
  Solving Inversion Problems in Atmospheric Sounding

  Working with Tony Gualtieri, Milt Halem, and George Serafino at NASA Goddard

- Dmitry A. Novik
  Data Compression, Image Analysis and Channel Efficient Coding

  Working with Jim Tilton & Manohar Mareboyana at NASA Goddard

- Elaine Finger, Johns Hopkins University
  Using Neural Networks for Machine Labeling of Satellite Data

  Working with Richard Kiang at NASA Goddard

- Leonard Dickens, University of Maryland
  Development of the SPUDS and IIFS System of the Intelligent Data Management Project

  Working with Nick Short at NASA Goddard

FELLOWSHIPS

FUNDED BY CRAY RESEARCH INC.

- 1991/92 Douglas Smith, Carnegie Mellon University
  (H.T. Kung, advisor)

  Intermediate Language and Virtual Architectures for High Performance Image Processing

- 1992/93 Kathleen Perez-Lopez, George Mason University
  (Arun Sood, advisor)

  Use of an Index/Browse Set of Images for Database Management

WORKSHOPS

- 1989 *Computing Challenges in Managing Future Massive Image Systems*

  Attended by 69 scientists and graduate students from 22 NASA, university, and private sector laboratories.

- 1990 *The Role of Computer Science Research in the Mission to Planet Earth*

  Attended by 238 scientists, graduate students, and business people.

- 1991 *Workshop on Parallel Algorithms (WOPA)*

  Provided monetary support for workshop held May 9 - 10 in New Orleans.

- 1992 *Scientific Data Management*

  To be held September 25. Expect 40-50 participants.

OTHER TASKS AND ACTIVITIES

- Technical Report Series
  - 86 titles currently.
  - Have filled requests for 1200+ copies thru July 31, 1992.

- Seminar Series
  Have organized and presented *Advances in Computational Sciences Seminar Series,* a joint CESDIS/ Space Data and Computing Division series, since September 1989.

- Data Compression Conference
  - DCC '91 April 7 - 11, 1991, Snowbird, Utah
    Provided publicity, registration, and on-site conference support.
  - DCC '92 March 24 - 26, 1992, Snowbird, Utah
    Provided publicity, registration, and on-site conference support.

- Peer Review Support for NASA New Research Announcements
  - 1990 Applied Information Systems Research
  - 1992 Research in High Performance Computing

OTHER TASKS AND ACTIVITIES

- Provide support to Minority University Space Interdisciplinary Network (MU-SPIN) project.
  - Provide half-time administrative support for Program Coordinator.
  - Supported annual Working Group Conference in September 1991.
  - Will support Working Group Conference in September 1992.
- Have provided access to work space and Sun computers for varying periods of time for 16 visitors to NASA Goddard since September 1989.

NEW CESDIS DIRECTIONS

- Hire full-time CESDIS Associate Director.
- Hire full-time CESDIS Staff Scientist.
- More coordination of AISRP Projects.
- Issue HPCC - CESDIS Call for Proposals.
- Assist Earth and Space Science HPCC projects.
- Build ties with ESDIS development (formerly EOSDIS).

BOTTLENECKS AND PROBLEM AREAS IN THE EARTH SCIENCES

Dr. Michael E. Botts
University of Alabama in Huntsville

August 13, 1992

Bottlenecks and Problem Areas In the Earth Sciences

AISRP Workshop II
Boulder, CO
August 11-13, 1992

Mike **Botts, Ph.D.**
Earth Systems Science Laboratory
University of Alabama in Huntsville

Earth Systems Science Laboratory
University of Alabama in Huntsville
Huntsville, AL 35899

Major General Issues

- Large Data Sets
- Lots of Data
- Interuse of Multiple Data Sets from Multiple Disciplines
- Temporal Data Important
- Multiband, Multiparameter Data

Earth Systems Science Laboratory
University of Alabama in Huntsville
Huntsville, AL 35899

Large Data Sets/Lots of Data

- Keeping Interactivity in Visualization
  - Adaptive Sampling Techniques could help
- Automatic Feature Recognition and Tracking
- Data Compression
  - Lossy vs Lossless - How much loss is acceptable
  - Analyzing/Visualizing Compressed Data without Decompression?

Earth Systems Science Laboratory
University of Alabama in Huntsville
Huntsville, AL 35899

Interuse of Multiple Data Sets from Multiple Disciplines

- Navigation, Gridding, Projection Issues
  - Common grids/projections corrupt data; no consensus
  - Leave navigation/gridding til as late as possible
  - How should navigation data be presented in files?
  - Should we use adaptive sampling grids?
  - How should we deal with irregularly-spaced, sparse data?
- Classification based on Multiple Data Sets
- Formats

Earth Systems Science Laboratory
University of Alabama in Huntsville
Huntsville, AL 35899

Navigation Data

- 3 ways:
  - provide location for each data point
  - use regular grid on standard projection and provide projection type, resolution, and tie points
  - leave data in low-level form (Level 2) and provide ephemeris data to require/allow tool to navigate
- Cycle: Types of tools available drives what provided in data format ... Data provided in format drives what we can do with tools
- Remember navigation includes: 3D space and time !!!

Earth Systems Science Laboratory
University of Alabama in Huntsville
Huntsville, AL 35899

Adaptive Sampling

- Advantages:
  - Keep details of data where variability high
  - Doesn't waste storage where variability low
  - Can aid interactivity in visualization and faster analysis of data
- Disadvantage:
  - requires availability of smart tools

Earth Systems Science Laboratory
University of Alabama in Huntsville
Huntsville, AL 35899

Temporal Elements

- Features and insight lost if not viewed within temporal frame
- Animation only partial answer
- Time series analysis
- Time averaging of asynchronous data
- Activity/event correlation

Earth Systems Science Laboratory
University of Alabama in Huntsville
Huntsville, AL 35899

Multiband/Multiparameter

- Classification schemes that use all the available data (e.g. hyperspectral data)
- Schemes based on sub-pixel treatment (e.g. spectral mixing)

Earth Systems Science Laboratory
University of Alabama in Huntsville
Huntsville, AL 35899

PART I:

THE STATE OF SCIENTIFIC VISUALIZATION WITH REGARD TO THE NASA EOS MISSION TO PLANET EARTH

PART II:

OPTIONS FOR MEETING THE VISUALIZATION REQUIREMENTS OF THE EOS MISSION

Mike Botts, Ph.D.
UAH

Earth Systems Science Laboratory
University of Alabama in Huntsville
Huntsville, AL 35899

COMPONENTS OF VISUALIZATION

- Image Processing
- 2D Data Plotting
- 3D Computer Graphics and Animation
- Volume Rendering
- Geographic Information Systems (GIS)
- Computer-Human Interactions (CHI)

Earth Systems Science Laboratory
University of Alabama in Huntsville
Huntsville, AL 35899

APPLICATIONS OF VISUALIZATION

- Scientific Investigation
- Data Validation
- Model and Algorithm Development and Validation
- Data Browse
- Information Transfer
- Mission Operations

Earth Systems Science Laboratory
University of Alabama in Huntsville
Huntsville, AL 35899

CHANGING ROLES AND CAPABILITIES

- CPU power increasing by order of magnitude every 3-4 years
- Transitions:
  - centralized to distributed computing
  - batch-mode operation to interactive computing
- Role of Computer Specialist :
  - moving from backend to frontend service

Earth Systems Science Laboratory
University of Alabama in Huntsville
Huntsville, AL 35899

"The most important immediate direction for visualization efforts is that of putting useful/usable (U^2), interactive tools into the hands of the scientist ."

Earth Systems Science Laboratory
University of Alabama in Huntsville
Huntsville, AL 35899

WHERE ARE WE NOW ?

- Great potential, not yet realized
- Powerful Components, Immature Tools
- Scattered, Incompatible Capabilities
- Scientists still overly dependent on computer specialists for tool use
- Scientists cannot yet take full advantage of the power of visualization on a routine basis

Earth Systems Science Laboratory
University of Alabama in Huntsville
Huntsville, AL 35899

WHY AREN'T SCIENTISTS USING WHAT'S AVAILABLE ?

- Tool not extensible or is too inflexible
- Tool too difficult
- Difficult to get existing data into tool
- Data & tools in complex heterogeneous computing environment
- Tool does not do what's needed
- Scientists not aware of tool

Earth Systems Science Laboratory
University of Alabama in Huntsville
Huntsville, AL 35899

WHY AREN'T SCIENTISTS USING WHAT'S AVAILABLE ? (cont)

- Tool too costly
- Scientist does not have access to adequate hardware for running tool
- Too difficult to communicate results
  - difficult to get adequate color prints or video
  - impossible to interact remotely with colleagues
  - publishing industry too archaic for color hardcopy, animation videos, algorithm and application exchange, and voice and sound annotation

Earth Systems Science Laboratory
University of Alabama in Huntsville
Huntsville, AL 35899

WHERE ARE THE VISUALIZATION BOTTLENECKS ?

- With few exceptions, hardware not major limitation of visualization environment
- Primary bottleneck is lack of adequate software allowing scientist to interactively visualize and analyze data on a routine basis

Earth Systems Science Laboratory
University of Alabama in Huntsville
Huntsville, AL 35899

COTS, PUBLIC DOMAIN, AND IN-HOUSE DEVELOPMENT

- Advantages of relying on COTS:
  - NASA pays only for working product, not development
  - Products tend to be more polished
  - Maintenance & support less of an issue
  - There is generally more awareness of product
- Disadvantages of COTS
  - Has traditionally been inflexible & not extensible
  - No access to code for porting or modification
  - Designed for large market appeal

Earth Systems Science Laboratory
University of Alabama in Huntsville
Huntsville, AL 35899

COTS, PUBLIC DOMAIN, AND IN-HOUSE DEVELOPMENT (cont)

- Disadvantages of COTS (cont)
  - Higher cost to user, but not necessarily NASA, as a whole
  - Users less apt to "try it out"
  - No direct control of directions for future
  - No assurance developer will be around next year
  - Development generally outside of scientists' realm
  - Approach for meeting needs generally more conservative and less creative

Earth Systems Science Laboratory
University of Alabama in Huntsville
Huntsville, AL 35899

COTS, PUBLIC DOMAIN, AND IN-HOUSE DEVELOPMENT (cont)

- Advantages of in-house development:
  - Designed to meet general & specific objectives
  - Source code available
  - Code adaptable & available for cannibalization
  - New creative & leading edge products often result
- Disadvantages of in-house development:
  - NASA assumes risk of development
  - Maintenance, distribution, and support difficult
  - Risk of costly dinosaurs

Earth Systems Science Laboratory
University of Alabama in Huntsville
Huntsville, AL 35899

COTS, PUBLIC DOMAIN, AND IN-HOUSE DEVELOPMENT (cont)

- Needs for in-house development:
  - Extension, modification, and integration of existing software to meet application-specific requirements
  - Contributions of useful modules or components to general programming pool
  - Leading edge (experimental) development of techniques and applications
  - Application-specific programs not available from COTS & not able to be created by extension of existing software

Earth Systems Science Laboratory
University of Alabama in Huntsville
Huntsville, AL 35899

COTS, PUBLIC DOMAIN, AND IN-HOUSE DEVELOPMENT (cont)

- Advantages & disadvantages of each
- Proper balance between these is advantageous
- Proper balance must be met with adequate plan and funding for each
- If NASA intends to rely more on COTS, we must improve interaction with vendors

Earth Systems Science Laboratory
University of Alabama in Huntsville
Huntsville, AL 35899

CURRENT DEVELOPMENT ENVIRONMENT WITHIN NASA

- Development in both OSSA and OAST have resulted in significant, leading-edge visualization and analysis tools
- Examples:

FAST	PLATO
Linkwinds	SPICE-lib / OOSPICE
IISS Image Spreadsheet	Imagic / Motif-view
Vis-AD	"orbital manuever"

Earth Systems Science Laboratory
University of Alabama in Huntsville
Huntsville, AL 35899

CURRENT DEVELOPMENT ENVIRONMENT WITHIN NASA (cont)

- Present development environment at OSSA can be characterized as:
  - Fragmented, with little direction or coordination
  - Lacking understandable structure for funding of both general and application-specific development
  - Lacking firm commitment or plan for meeting true visualization needs of EOS science community
  - Lacking adequate mechanisms for technology transfer
  - Experiencing inefficient use of limited funding

Earth Systems Science Laboratory
University of Alabama in Huntsville
Huntsville, AL 35899

DIAMONDS AND DINOSAURS

- Diamonds = experimental development efforts which hold much promise for scientists or developers
- Dinosaurs = old development programs with limited momentum & decreasing application

Earth Systems Science Laboratory
University of Alabama in Huntsville
Huntsville, AL 35899

THE CHALLENGE WITH DIAMONDS

- What do you do once they become successful ?
- Without proper nurturing, diamonds quickly falter due to stagnation because of overwhelming requirements of user support
- Need adequate mechanisms for transitioning diamonds to maturity or to industry

Earth Systems Science Laboratory
University of Alabama in Huntsville
Huntsville, AL 35899

CHALLENGE WITH DINOSAURS

- What do you do with archaic programs still in use ?
- Often offer more functionality than other existing programs
- Archaic with regard to:
  - interfaces, program structure, data formats, graphics capabilities, and portability to modern platforms
- Porting to new machines is *not* adequate

Earth Systems Science Laboratory
University of Alabama in Huntsville
Huntsville, AL 35899

HELP PREVENT FUTURE DINOSAURS

- Flexibility and extensibility
- Object-oriented principles of abstraction and modularity
- Assist transition to maturity and beyond
- Have mechanisms in place for graceful phasing out of program

Earth Systems Science Laboratory
University of Alabama in Huntsville
Huntsville, AL 35899

MAIN AREAS FOR CONSIDERATION

- Integration of Visualization with Data Management and Analysis
- Application Development for the Scientist
- Providing Extensibility without Complexity
- Application-Specific Development without Redundant Programming
- Simplification of Complex Heterogeneous Computing Environment

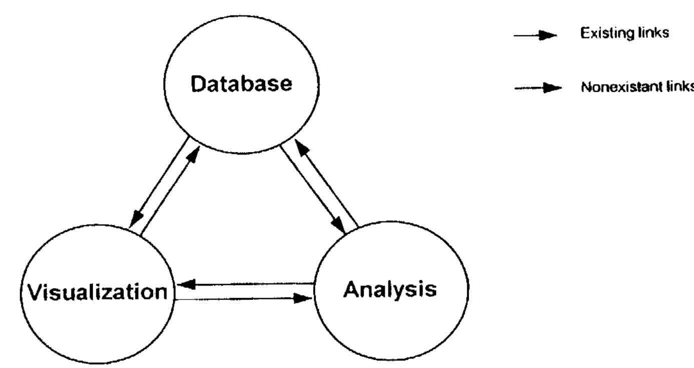
Earth Systems Science Laboratory
University of Alabama in Huntsville
Huntsville, AL 35899

MAIN AREAS FOR CONSIDERATION (cont)

- Development of New Techniques and Components
- Education and Communication
- Distribution and Maintenance

Earth Systems Science Laboratory
University of Alabama in Huntsville
Huntsville, AL 35899

INTEGRATION OF VISUALIZATION WITH DATA MANAGEMENT AND ANALYSIS



Earth Systems Science Laboratory
University of Alabama in Huntsville
Huntsville, AL 35899

MAIN AREAS FOR CONSIDERATION

- Integration of Visualization with Data Management and Analysis
  - Visualization - Analysis link is a must!
  - Visualization/Analysis - Data Management important, but less time critical
  - Navigation is subset of this issue that spans all components of triad

Earth Systems Science Laboratory
University of Alabama in Huntsville
Huntsville, AL 35899

MAIN AREAS FOR CONSIDERATION

- Application Development for Scientists as End-Users
  - Do what they need done !
  - Ease of use
  - Interfaces and structure intuitive to scientist
  - Allow complexities to be hidden

Earth Systems Science Laboratory
University of Alabama in Huntsville
Huntsville, AL 35899

MAIN AREAS FOR CONSIDERATION

- Providing Extensibility Without Complexity
  - Extensibility critical to success of COTS
    - » examples: AVS, SGI Explorer, Khoros, IDL, PV-Wave
  - At present, extensibility = complexity
  - Allow hiding of complexities

Earth Systems Science Laboratory
University of Alabama in Huntsville
Huntsville, AL 35899

MAIN AREAS FOR CONSIDERATION

- Application-Specific Programming Without Redundant Programming
  - Redundant programming is extensive
  - Availability of more extensible and flexible tools is helping
  - Application-specific programming should = module building when possible
  - Modularity and abstraction are critical

Earth Systems Science Laboratory
University of Alabama in Huntsville
Huntsville, AL 35899

MAIN AREAS FOR CONSIDERATION

- Simplification of Complex Heterogeneous Computing Environment
  - Sources:
    - » Operating systems, windowing systems, graphics devices, data formats, network protocols, output devices
  - Issues consume estimated 60% of efforts
  - Homogenization only partial answer
  - Hiding complexities from scientists is crucial

Earth Systems Science Laboratory
University of Alabama in Huntsville
Huntsville, AL 35899

MAIN AREAS FOR CONSIDERATION

- Development of New Techniques and Components
  - In few cases, involves new approaches and thus requires complete application development
    - » examples: IISS, Linkwinds, Vis-AD
  - In most cases, should only require development of module or component for existing applications

Earth Systems Science Laboratory
University of Alabama in Huntsville
Huntsville, AL 35899

MAIN AREAS FOR CONSIDERATION

- Education and Communication
  - Scientists lack understanding of capabilities and limitations of visualization
  - Project managers unsure if needs are being met
  - Developers lack understanding of true needs
  - Scientists unaware of existence and capabilities of many tools

Earth Systems Science Laboratory
University of Alabama in Huntsville
Huntsville, AL 35899

MAIN AREAS FOR CONSIDERATION

- Distribution and Maintenance
  - In-house: transition, distribution, maintenance, and user support are major issues
  - COTS: licensing and "trying out" in distributed environment is challenging
  - Public Domain: NASA should support useful public domain tools

Earth Systems Science Laboratory
University of Alabama in Huntsville
Huntsville, AL 35899

PART I: CONCLUSIONS

- Visualization vital for meeting scientific objectives of EOS Mission
- Our present suite of visualization techniques is impressive
- However, programs capable of putting this power into the hands of scientists is lacking
- COTS software can probably serve greater role in future, but still need for in-house development, as well

Earth Systems Science Laboratory
University of Alabama in Huntsville
Huntsville, AL 35899

OBJECTIVES

- Increase effective use of visualization / analysis tools by scientists
- Maximize the return on development efforts

Earth Systems Science Laboratory
University of Alabama in Huntsville
Huntsville, AL 35899

IDEAS FOR MEETING VISUALIZATION REQUIREMENTS

- Establish a Visualization/Analysis Initiative
- Visualization/Analysis Working Group
- Visualization/Analysis Assistance "Centers"
- Streamline NASA-funded Development
- Use Pathfinders to Test Full Data Path
- Set up Vendor Programs
- Improve Licensing/Procurement
- Improve publishing and Remote Interaction

Earth Systems Science Laboratory
University of Alabama in Huntsville
Huntsville, AL 35899

IDEAS FOR MEETING VISUALIZATION REQUIREMENTS

- Establish a Visualization/Analysis Initiative
  - To assure the availability of visualization and analysis tools to meet EOS objectives
  - To work toward integration of Visualization, Analysis, and Data Management
  - To foster the successful application of COTS and public domain tools
  - To establish workable plans for meeting in-house development needs

Earth Systems Science Laboratory
University of Alabama in Huntsville
Huntsville, AL 35899

IDEAS FOR MEETING VISUALIZATION REQUIREMENTS

- Visualization/Analysis Working Group
  - tool-based objectives
  - mix of earth scientists and computer scientists, mix of visionaries and practical minded
  - To remain aware of the state of visualization as it relates to the needs of EOS scientists
  - To identify areas of deficiencies in the present suite of visualization/analysis tools
  - To recommend options for alleviating any deficiencies which might exist

Earth Systems Science Laboratory
University of Alabama in Huntsville
Huntsville, AL 35899

IDEAS FOR MEETING VISUALIZATION REQUIREMENTS

- Visualization/Analysis Assistance "Centers"
  - located at one or more DAACs
  - service-based (by phone, network, or visits)
  - To provide knowledgeable assistance to scientists regarding availability, capability, and use of tools
  - To promote education and awareness of tools
  - To serve as liason to vendors regarding scientists needs and existing gaps
  - To recommend or perform small-scale development to integrate or extend existing tools to fill needs

Earth Systems Science Laboratory
University of Alabama in Huntsville
Huntsville, AL 35899

IDEAS FOR MEETING VISUALIZATION REQUIREMENTS

- Streamline NASA-funded Development
  - Assure availability of low-level libraries & encourage use
  - Move NASA into object oriented programming (by education & assistance, not by force)
  - Set up apropriate and understandable funding structure for both general and application-specific development
  - Encourage development of modules for application-specific needs, where possible
  - Set up mechanisms for transition of diamonds and for replacing dinosaurs

Earth Systems Science Laboratory
University of Alabama in Huntsville
Huntsville, AL 35899

IDEAS FOR MEETING VISUALIZATION REQUIREMENTS

- Low-Level Libraries
  - Data Input
  - Data Analysis/Statistics/Contouring
  - Navigation/Coregistration/Projections
  - Image Processing
  - Visualization (2D plotting, 3D surfaces, etc)
  - Output to Video and Print
  - Widgets (controls, histogram displays, etc.)
  - Distributed Processing

Earth Systems Science Laboratory
University of Alabama in Huntsville
Huntsville, AL 35899

IDEAS FOR MEETING VISUALIZATION REQUIREMENTS

- **Use Pathfinders to Test Full Data Path**
  - Are the tools available to allow interuse of data for real science?
  - What requirements do the tools place on the data sets?
  - What is the proper approach: smart tools or common data?
  - Navigation of data within tools, and as late as possible, is critical

Earth Systems Science Laboratory
University of Alabama in Huntsville
Huntsville, AL 35899

IDEAS FOR MEETING VISUALIZATION REQUIREMENTS

- **Set up Vendor Programs**
  - Include software vendors (developers) in workshops
  - Set-up vendor-on-site programs
  - Consider SBIR-like funding programs for vendors to extend existing tools to meet EOS objectives
  - Consider programs to transition in-house diamonds to commercial industry

Earth Systems Science Laboratory
University of Alabama in Huntsville
Huntsville, AL 35899

IDEAS FOR MEETING VISUALIZATION REQUIREMENTS

- **Improve Licensing and Procurement**
  - Work with Vendors to encourage:
    - » "try-it-out" distribution copies
    - » floating licenses
    - » "pay-as-you-use" licenses
  - Speed up procurement of hardware and software

Earth Systems Science Laboratory
University of Alabama in Huntsville
Huntsville, AL 35899

IDEAS FOR MEETING VISUALIZATION REQUIREMENTS

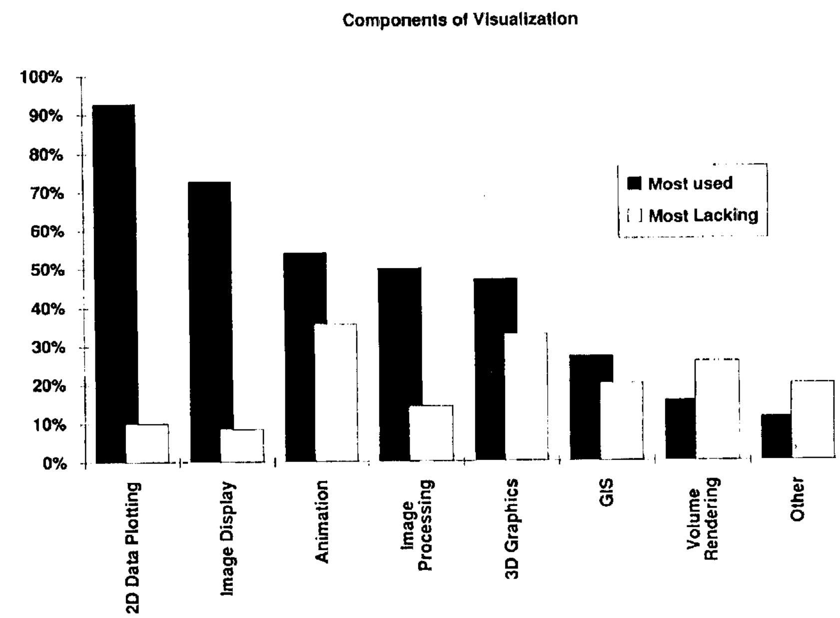
- **Improve publishing and Remote Interaction**
  - Encourage publishing industry to move into multi-media publishing (allow text, images, animation, binary and ascii programs, and sound)
  - Encourage the development of tools allowing remote interaction between scientists
  - Assure availability of multimedia exchange capabilities among EOS scientists, as well as EOS project personnel

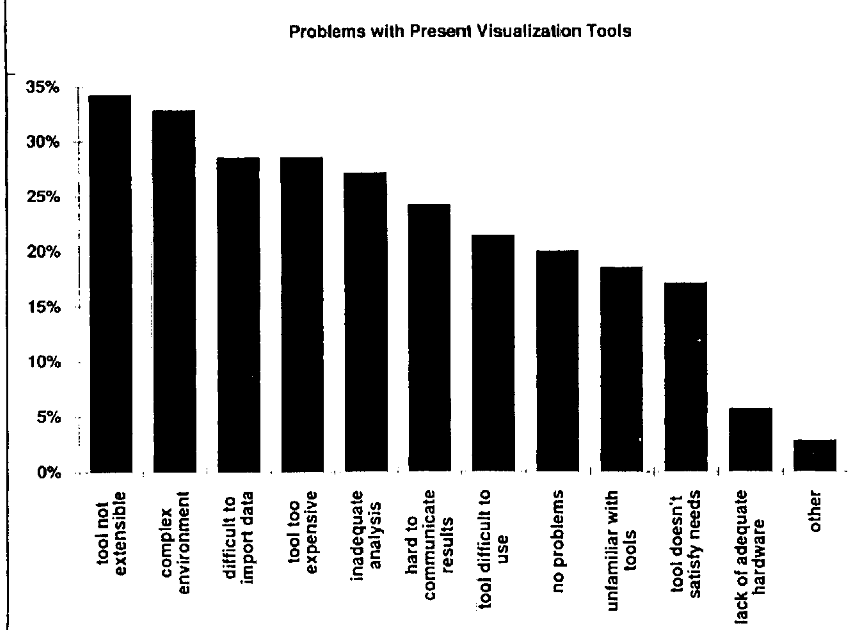
Earth Systems Science Laboratory
University of Alabama in Huntsville
Huntsville, AL 35899

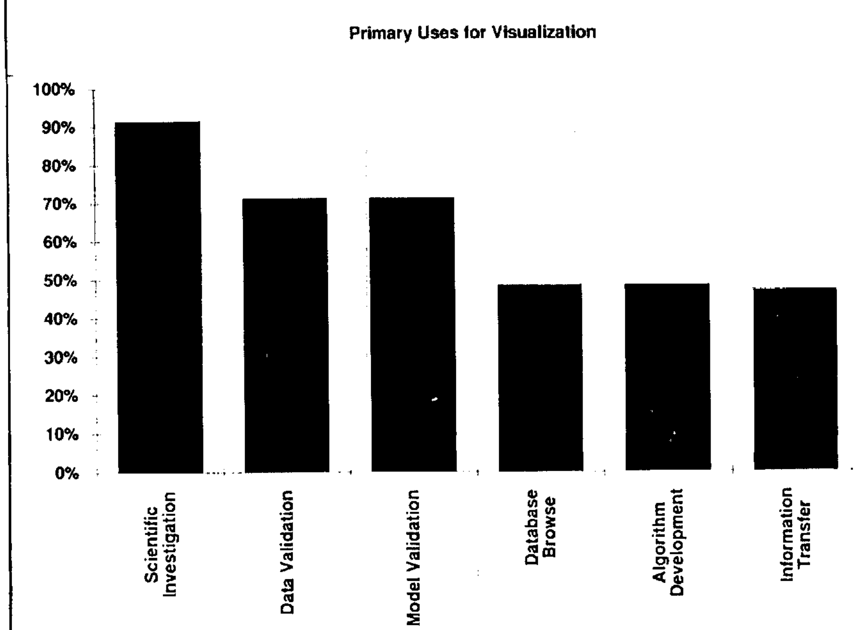
CONCLUSIONS

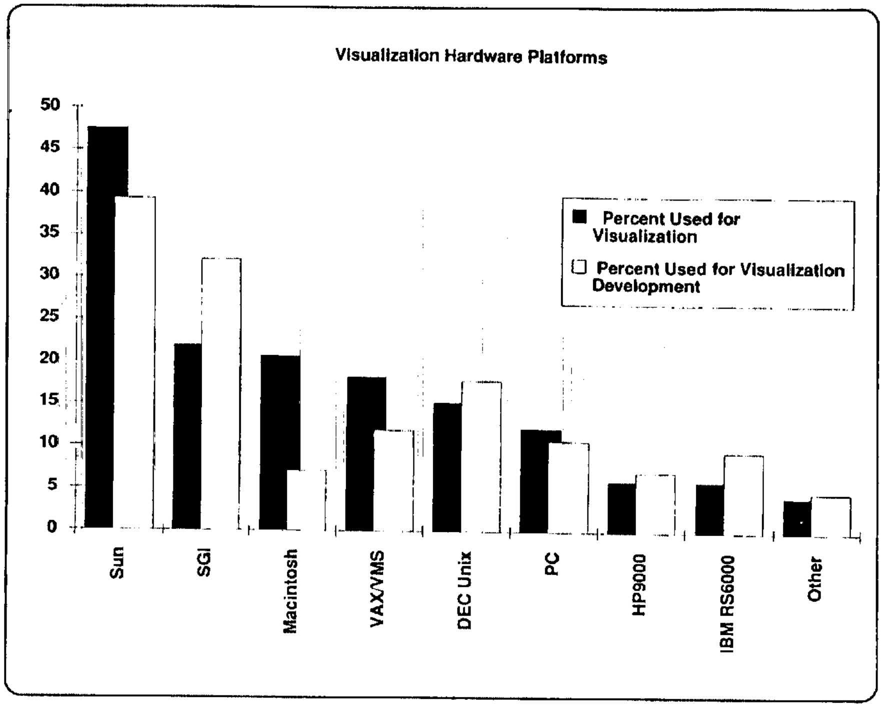
- Visualization is a vital link required to meet the scientific objectives of EOS
- Although there is a lot of power in our present techniques, the tools are not yet adequate
- The answers do not involve much greater investments of NASA funding
- The answers do involve having visions for meeting needs, and appropriate programs to balance COTS and in-house development

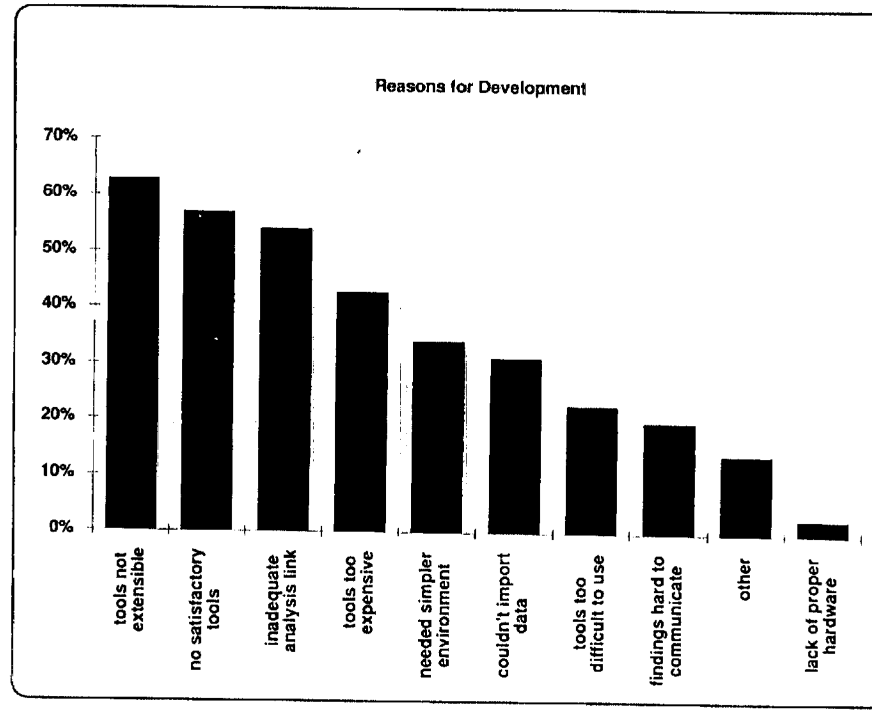
Earth Systems Science Laboratory
University of Alabama in Huntsville
Huntsville, AL 35899

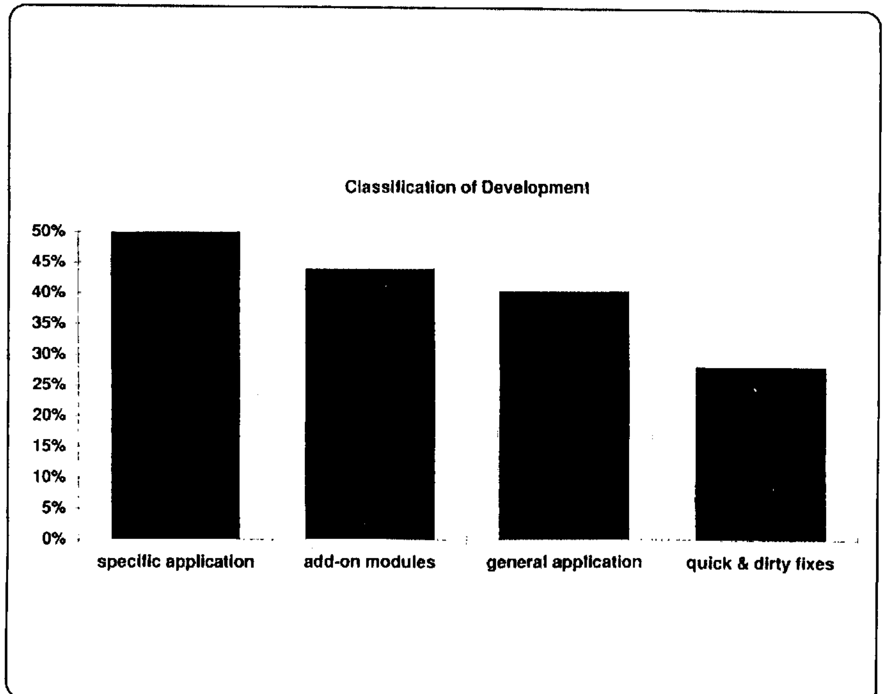


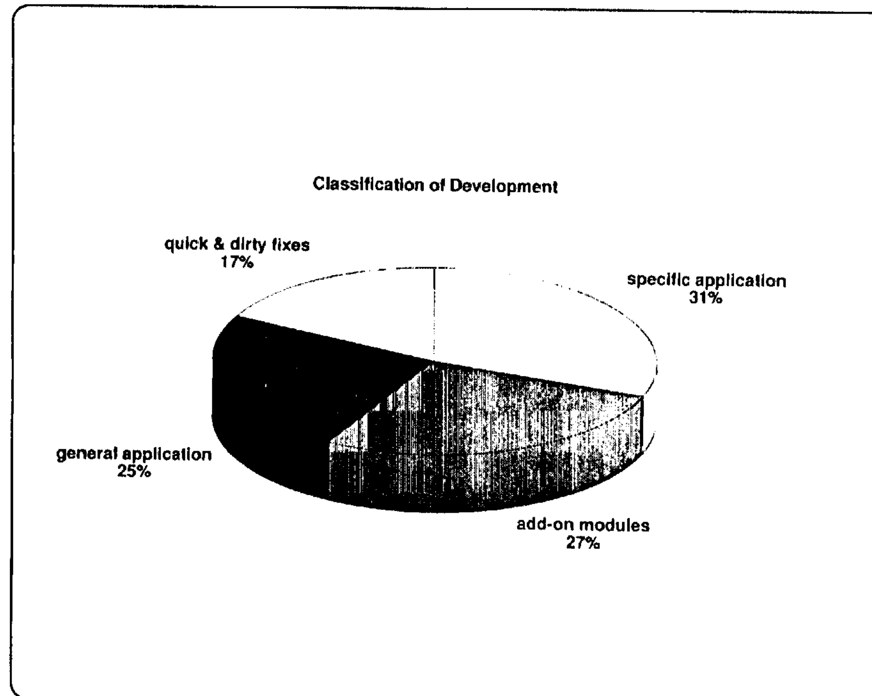





Visualization Hardware Platforms
50
45
40
35
30
25
20
15
10
5
0
Percent Used for Visualization
Percent Used for Visualization Development
Sun
SGI
Macintosh
VAX/VMS
DEC Unix
PC
HP9000
IBM RS6000
Other


Reasons for Development
70%
60%
50%
40%
30%
20%
10%
0%
tools not extensible
no satisfactory tools
inadequate analysis link
tools too expensive
needed simpler environment
couldn't import data
tools too difficult to use
findings hard to communicate
other
lack of proper hardware


Classification of Development
50%
45%
40%
35%
30%
25%
20%
15%
10%
5%
0%
specific application
add-on modules
general application
quick & dirty fixes


Classification of Development
quick & dirty fixes 17%
specific application 31%
general application 25%
add-on modules 27%

OSSA LIFE SCIENCES DIVISION RESEARCH REQUIREMENTS FOR SPACE STATION FREEDOM

Dr. Robert Jackson
Ames Research Center

August 13, 1992

OSSA LIFE SCIENCES DIVISION RESEARCH REQUIREMENTS for SPACE STATION FREEDOM

Robert Jackson
Centrifuge Facility Project Office
NASA Ames Research Center
Moffett Field, CA 94306

OUTLINE

* Life Sciences Program for Space Station Freedom
* Research Scenarios
  - Descriptions
  - Data Flows
  - Bottlenecks and Research Needs
* Summary

SPACE STATION FREEDOM PROGRAM

Centrifuge Facility (CF)
A suite of equipment to support research examining the influence of microgravity and radiation on biological processes, and to test artificial gravity as a countermeasure.

Gravitational Biology Facility (GBF)
A suite of general-use life sciences laboratory equipment that will support investigations of cell, developmental, and plant biology. The GBF equipment will be composed of existing space-qualified or off-the-shelf equipment as much as possible.

Controlled Ecological Life Support System Test Facility (CTF)
Equipment to serve as a testbed for studying crop growth and productivity in microgravity.

PROGRAM (Continued)

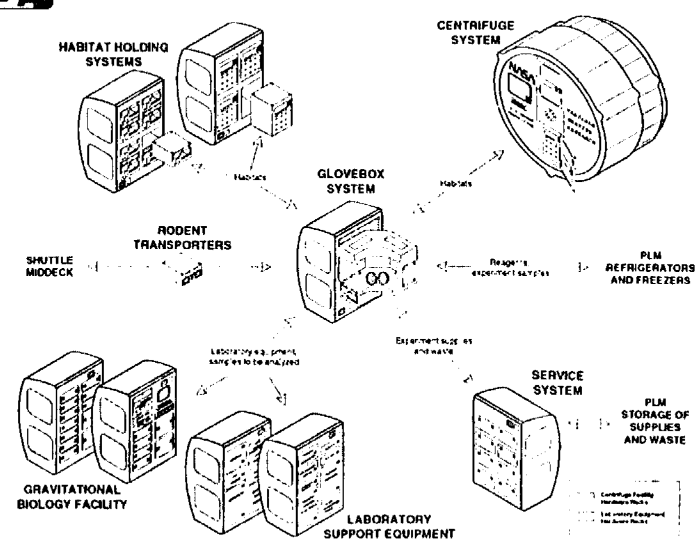
Gas-Grain Simulation Facility (GGSF)
Equipment to investigate fundamental chemical and physical interactions of suspended particles and to measure the chemical and physical properties of the resulting materials

Biomedical Monitoring and Countermeasures Facility (BMAC)
Equipment to monitor and maintain crew health and performance during long duration missions.

PROGRAM STATUS

Facility	Number of Racks	Development Status
CF	4 + Centrifuge	Phase C/D RFP released 1/1/93 contract start
GBF	2	Phase A study
CTF	2 adjacent	Phase A study
GGSF	1	Phase A study
BMAC	4	Definition phase

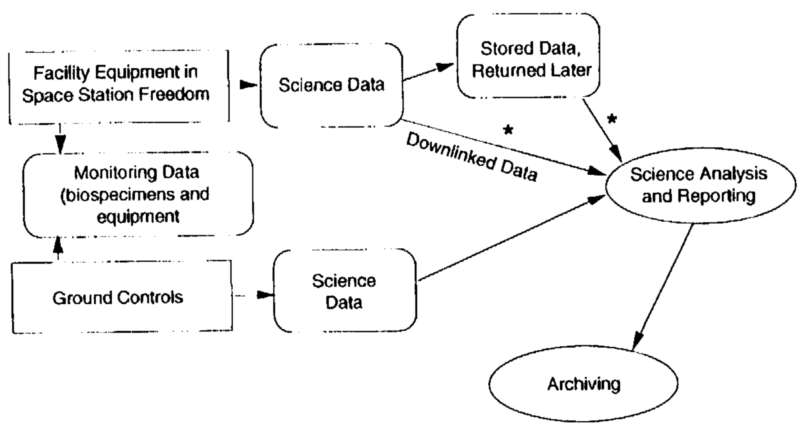
CENTRIFUGE FACILITY EQUIPMENT



SCENARIO - AUTOMATIC OPERATION

- Mode provides routine operation of major facility equipment to support biospecimen growth and development
- Mode features
  - Equipment operates automatically according to programmed plan
  - Planning, monitoring, command, and control from ground operations center
  - Coordinated ground control experiments pre-flight, simultaneous, post-flight
  - Science data downlinked and distributed to PIs

AUTOMATIC OPERATION DATA FLOW



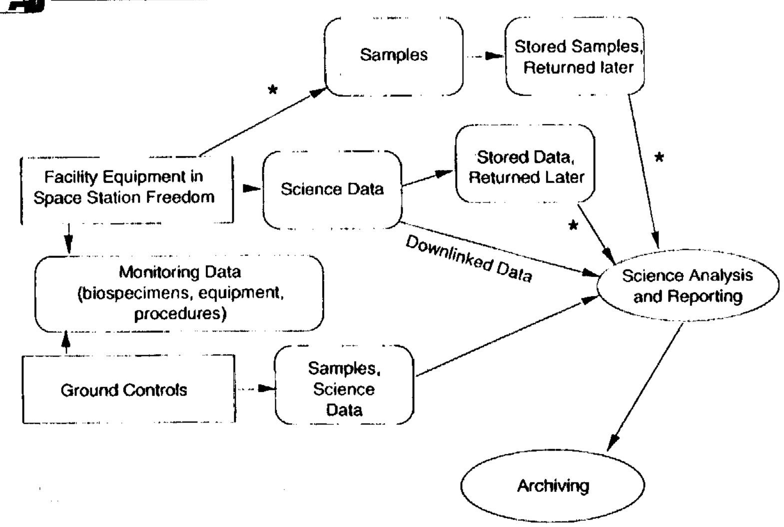
AUTOMATIC OPERATION STATUS

- Bottlenecks
  - Limited downlink bandwidth
  - Downlink interruptions
  - No on-board communications outage recorder
- Some research needs
  - Tailored data bandwidth reduction techniques
  - Efficient on-board data processors and storage

SCENARIO - CREW OPERATED

* Mode provides crew tended operation to set up runs, manipulate specimens, collect data, collect samples, analyze samples, maintain health, maintain equipment, etc
- Mode features
  - Crew conducts activities during scheduled periods according to procedures
  - Crew interacts with the ground for advice, checking, reporting, procedure modifications
  - Coordinated ground control experiments pre-flight, simultaneous, post-flight
  - Some data downlinked and distributed to PIs
  - Specimens, samples, remaining data returned later and distributed to PIs

CREW OPERATION DATA FLOW



CREW OPERATION STATUS

- Bottlenecks
  - Crew time
  - Crew skill maintenance
  - Limited uplink and downlink bandwidths
  - Uplink and downlink interruptions
  - Late delivery of data and samples
  - Long storage time of samples

- Some research needs
  - Effective on-board training aids
  - Tailored data bandwidth reduction techniques
  - Efficient on-board data processors and storage

SUMMARY

- Life Sciences program for Space Station Freedom is proceeding
  - Initial science objectives have been defined
  - Equipment is being developed

- Research is needed to enable effective conduct of advanced experiments
  - Accommodate limited uplink and downlink
  - Minimize demand for power and volume
  - Maintain and improve crew skills

ARCHIVING PLANETARY MISSION DATA

Dr. Steven Lee
Laboratory for Atmospheric and Space Physics
University of Colorado

August 13, 1992

ARCHIVING PLANETARY MISSION DATA

Steve Lee
Univ. Colorado / LASP
PDS Atmospheres Discipline Node

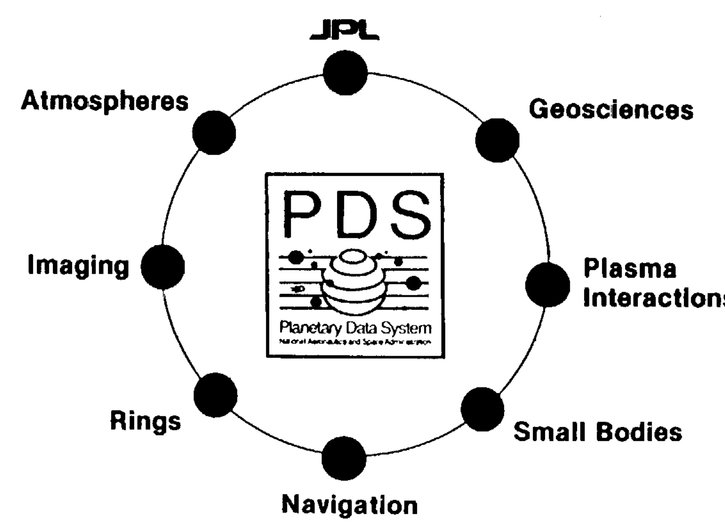
THE PLANETARY DATA SYSTEM

INTENT:
- Preserve science results of planetary exploration missions
- Provide ready access of data to research community
- Promote analysis of planetary data

STRUCTURE:
- Central Node
- Discipline Nodes
- Data Nodes
- National Space Science Data Center (NSSDC)

PDS DATA PRESERVATION / DISTRIBUTION
- **Preserve data sets**
  - assemble documentation
  - assemble data
  - apply standards
  - validation
  - peer review
- **Curate data sets**
- **Distribute data sets**
- **Provide scientific expertise**
- **Promote research**

Planetary Data System



The Planetary Data System
Planetary Atmospheres Discipline Node

The Atmospheres Node is a consortium of institutions encompassing a broad range of scientific interests. The locations, areas of expertise, and personnel involved are:

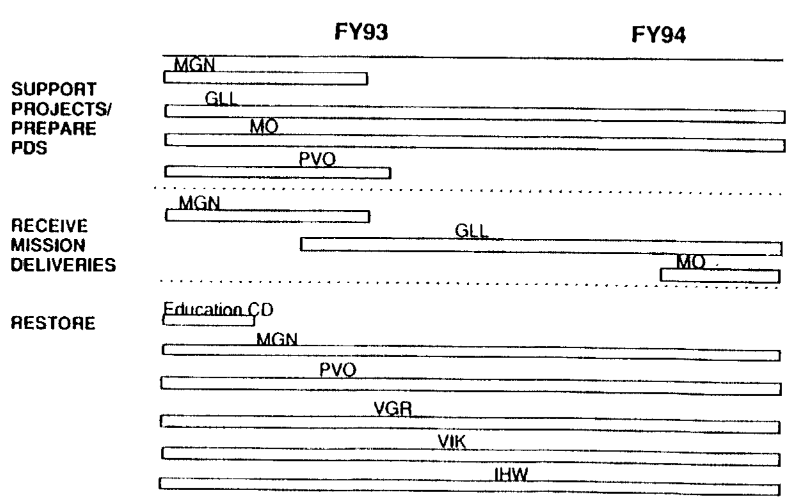
- **University of Colorado Discipline Node**
  UV spectroscopy, aeronomy, radiative transfer, microwave spectroscopy, surface-atmosphere interactions, cometary atmospheres, data management *(Manager: Steve Lee)*
- **Ames Research Center Subnode**
  Atmospheric dynamics, atmospheric modelling *(Manager: Bob Haberle)*
- **University of Arizona Subnode**
  UV, EUV spectroscopy *(Manager: Bill Sandel)*
- **Goddard Space Flight Center Subnode**
  IR radiometry, IR spectroscopy *(Manager: Barney Conrath)*
- **Jet Propulsion Laboratory Subnode**
  Spectral modelling, radiative transfer, radio science, dynamics, chemistry, molecular spectroscopy *(Manager: Glenn Orton)*
- **New Mexico State University Subnode**
  Imaging of planetary atmospheres, climatic data from long-term monitoring *(Manager: Reta Beebe)*
- **University of Washington Subnode**
  Meteorology, atmospheric dynamics, modelling *(Manager: Conway Leovy)*

Archiving Planetary Data
Active Missions

PROBLEM AREAS

- Obligation of flight projects to archive data
- Preservation of mission funds for data archiving
- Availability of project personnel for archiving tasks
- Definition of interfaces between projects and PDS
- Large lead times needed to influence archiving plans
- Selection of storage technology to ease archiving
- Massive data volumes (hundreds of CDROM volumes)
- Proprietary period and scheduling of data transfer (mission —> PDS)

PDS MISSION ACTIVITIES (FY93 - FY94)



TECHNOLOGY TRANSFER

Dr. Tom Handley
Dr. Larry Preheim
Jet Propulsion Laboratory

August 13, 1992

TECHNOLOGY TRANSFER

Tom Handley
Larry Preheim

JPL

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

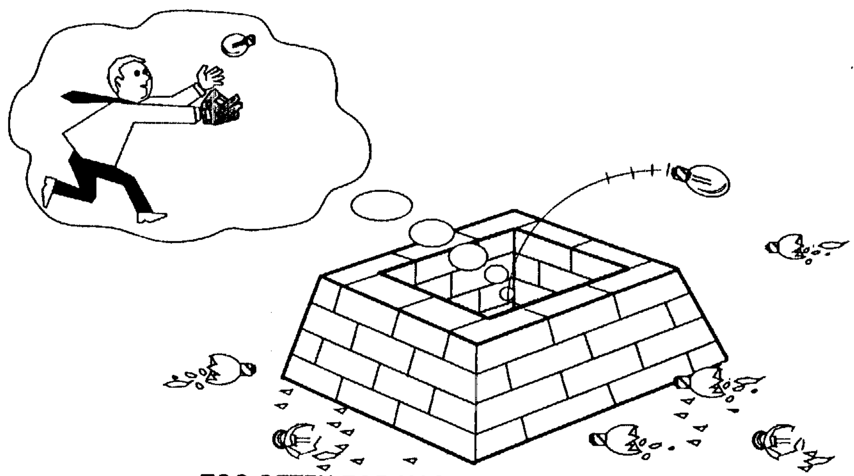
JPL TECHNOLOGY TRANSFER PURPOSE AND OUTCOME

- PURPOSE: RAISE AWARENESS OF THE PROBLEM AND EDUCATE ABOUT THE PROCESS
- SUGGESTED OUTCOME: WE AGREE ON A USABLE SET OF METHODOLOGIES FOR IMPLEMENTING TECHNOLOGY TRANSFER

JPL TECHNOLOGY TRANSFER DEFINITION

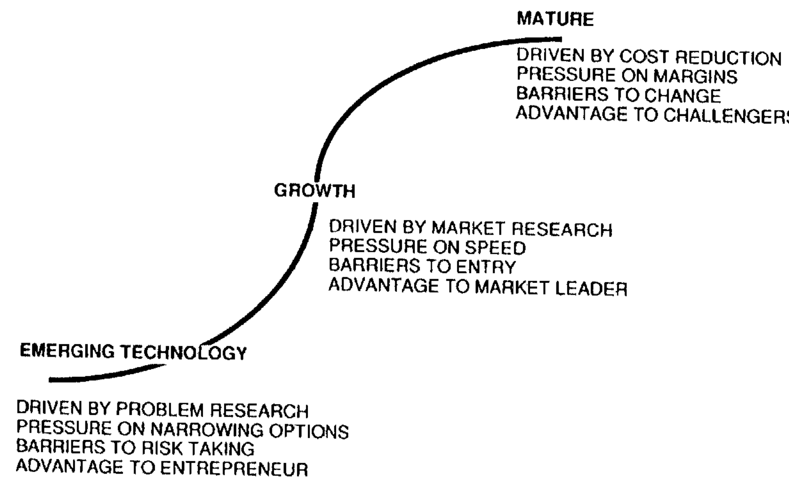
- THE TRANSFER OF ORGANIZED KNOWLEDGE TO A PROJECT OR PROGRAM FOR THE EVENTUAL PURPOSE OF PRODUCING NEW OR IMPROVED, PRODUCTS, PROCESSES OR SERVICES.
- TRANSFER WILL OCCUR THROUGH ONE, OR MORE, OF THE FOLLOWING MODES:
  - OCCASIONAL CONSULTING
  - DOCUMENTATION (REPORTS, ASSESSMENTS, PROGRAMS, OR DRAWINGS)
  - TRAINING (ON-THE-JOB, ON-SITE OR ELSEWHERE)
  - DEMONSTRATION (PROOF-OF-PRINCIPLE OR APPLICATION TO A REAL-WORLD PROBLEM)
  - COLLABORATIVE TECHNICAL WORK.

JPL TRADITIONAL TECHNOLOGY TRANSFER



TOO OFTEN R&D HAS BEEN CONTENT TO 'THROW ITS PRODUCT OVER THE WALL AND HOPE SOMEONE WILL CATCH IT.'

JPL IMPLICATIONS OF TECHNOLOGY MATURITY



JPL SIMPLIFIED LOOK AT BOTH SIDES

ISSUE	TECHNOLOGY OR ADVANCED DEVELOPMENT	IMPLEMENTATION OR PRODUCTION
· MANAGEMENT	· TECHNICALLY ORIENTED	· PRODUCT ORIENTED
· STAFFING	· TECHNOLOGIST AND SPECIALISTS	· ENGINEERS AND PRODUCTION PERSONNEL
· THROUGHPUT	· SMALL	· LARGE
· INERTIA	· LOW	· HIGH
· DOCUMENTATION	· MINIMAL	· EXTENSIVE
· COST	· NOT PRIMARY	· PRIMARY
· SUPPORT	· SMALL RESOURCE REQUIREMENT	· LARGE RESOURCE REQUIREMENT

JPL BARRIERS

- THE USER COMMUNITY LACKS A PROCESS TO IDENTIFY COMMON TECHNOLOGY REQUIREMENTS
- THE USER COMMUNITY LACKS A VEHICLE TO EXERT THE COLLECTIVE LEVERAGE TO CAUSE NASA TO IMPLEMENT COMMON DESIGN.
- RESOURCES INVESTED IN EXISTING SYSTEMS AND APPLICATIONS, AND THE ATTITUDE AND CULTURE OF THE WORK FORCE MAKE IT DIFFICULT TO EVOLVE TO NEW TECHNOLOGIES
- CURRENT PRACTICES ENCOURAGE A TACTICAL APPROACH TO SOLVING TECHNICAL PROBLEMS WHILE IGNORING KEY STRATEGIC (I.E. LONG TERM) ISSUES.
- THERE ARE INADEQUATE INCENTIVES FOSTERING THE INSERTION OF NEW TECHNOLOGY INTO NEW MISSIONS. THE LINKAGE BETWEEN TECHNOLOGY PAYBACK AND ACHIEVING MISSIONS GOALS IS NOT STRONG.
- FEAR OF BEING UNABLE TO COMPLETE A MISSION (ON-TIME, WITHIN BUDGET, AND MEETING MISSION GOALS) USING "NEWER" TECHNOLOGY.
- THERE IS NO DOCUMENTED, COHERENT NASA VISION FOR BROAD-BASED TECHNOLOGY INTEGRATION AND THE ROLE OF TECHNOLOGY TRANSFER IN ACHIEVING THAT VISION.

JPL

BARRIERS (cont'd)

- THERE IS NO SHARED VISION FOR DEVELOPING A TECHNOLOGY TRANSFER PROCESS.
- TRANSFER IS FURTHER COMPLICATED BY THE FACT THAT AFT TIMES CAPABILITIES RATHER THAN SPECIFIC PRODUCTS MUST BE TRANSFERRED.
- WITH TODAY'S PROJECTS, YOU CANNOT SIMULTANEOUSLY ACCEPT A "FIXED-PRICED" CONTRACT FROM CONGRESS TO DEVELOP A MAJOR UNDERTAKING AND AT THE SAME TIME SUPPORT TECHNOLOGY DEVELOPMENT AND THE UNAVOIDABLE ATTENDANT RISKS, I.E. COST UNCERTAINTY.
- INADEQUATE STAFFING BY ENGINEERING. A COMMON RESPONSE TO THE SUGGESTION FOR NEW TECHNOLOGY IS "WE DO NOT HAVE ANYONE HERE WHO HAS THE TECHNICAL SKILLS AND KNOWLEDGE TO INCORPORATE THIS TECHNOLOGY INTO CURRENT PROJECTS.
- THE PERCEPTION THAT A TECHNOLOGY IS TOO COMPLEX WILL OFTEN LEAD THE INTENDED USERS TO QUESTION THE TECHNOLOGY DEVELOPERS CREDIBILITY.
- NASA DOES NOT DEVELOP SERIOUS PLANS BEYOND A FIVE YEAR NEW START HORIZON

JPL

KEY T*T MAXIM

T*T IS A CONTACT SPORT

- OCCURS IN CONTEXT OF ONE-ON-ONE RELATIONSHIPS OF TECHNOLOGISTS/ORGANIZATIONS
- THE PROGRAMMATIC (MARKETING) CHALLENGE IS TO ESTABLISH THESE RELATIONSHIPS

JPL

CONVERTING PUSH TO PULL

AN EFFECTIVE MARKETING (OUTREACH?) PROGRAM <u>CONVERTS</u> T*T FROM PUSH TO PULL

- ONCE A USER (APPLIER) IS IDENTIFIED, HIS NEED DIRECTS THE R&D

IF ALL R&D IS DIRECTED

- SCIENTIFIC
- APPLICATIONS

THEN

SUPERIMPOSING SECONDARY R&D DIRECTIONS ON PRIMARY DIRECTIONS <u>CONFIRMS</u> T*T IS CONVERTED FROM PUSH TO PULL

JPL

MARKETING MODEL OF T*T

"...A DIRECTOR OF LICENSING FOR A "FORTUNE 100" MULTINATIONAL CORPORATION OBSERVED THAT THEY LONG AGO CONCLUDED THAT DISSEMINATION (OF INFORMATION) DID NOT PRODUCE RESULTS. HE MAINTAINED THAT THE ONLY SURE WAY TO TRANSFER (LICENSE) COMPANY DEVELOPED TECHNOLOGIES WAS TO MARKET, OR SELL, THEM IN THE SAME WAY ANY OTHER COMMERCIAL PRODUCT IS SOLD. FEDERAL AGENCY PROGRAMS HAVE NOT GONE, OR EVEN PLAN TO GO, THAT FAR. INDEED, CHANCES ARE THAT MOST FEDERAL AGENCIES DO NOT NOW HAVE EVEN A FAIR IN-HOUSE CAPABILITY TO DETERMINE THE POTENTIAL COMMERCIAL VALUES OF THEIR OWN TECHNOLOGIES."

KOONS, M.E., "THE TRANSFER OF TECHNOLOGY THROUGH AN INDUSTRIAL COOPERATION PROGRAM," *AMERICAN SOCIETY OF MECHANICAL ENGINEERS*, PP. 6, NEW YORK, 1975, QUOTED IN

KATHLEEN CONNELL, ET AL, *TECHNOLOGY TRANSFER IN THE NASA AMES ADVANCED LIFE SUPPORT DIVISION*, NASA AMES RESEARCH CENTER DOCUMENT NO.: 92-SAS-R-004, MARCH 17-19, 1992

JPL

KEY FACTORS

- PLANNING
- USER INVOLVEMENT
- COMMUNICATIONS
- A PROCESS IS REQUIRED
- KNOWING AND ASKING THE RIGHT QUESTIONS
- RESPONSIBILITY AND ACCOUNTABILITY
- FUNDING

JPL

POTENTIAL TECHNOLOGY TRANSFER SUPPORT SERVICES

1 FIRST LINE OF CONSULTING FOR INSTALLATION

2 DUPLICATE AND DISTRIBUTE THE TECHNOLOGY

3 SUPPORT CONFIGURATION MANAGEMENT

4 PROVIDE DEMONSTRATIONS OF THE TECHNOLOGY IN REAL WORLD SITUATIONS

5 ACTIVELY SELL THE TECHNOLOGIES

6 ASSIST IN PROBLEM IDENTIFICATION/RESOLUTION

7 CAPTURE RELIABILITY, ADAPTABILITY DOCUMENTATION HISTORY AND PROVIDE LESSONS LEARNED

8 COSMIC SUBMITTAL

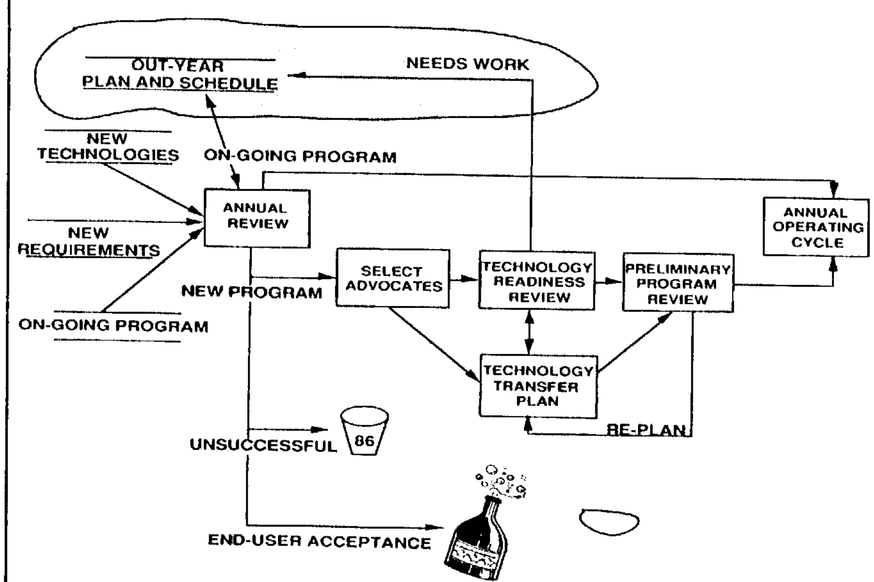
JPL

A DEVELOPERS PERSPECTIVE ON APPROACHES TO TECHNOLOGY TRANSFER

APPROACH	COMMITMENT	COST	CHARACTERISTICS/COMMENTS
COSMIC	LOW	LOW	MIXED EFFECTIVENESS
BULLETIN BOARDS	↓	↓	GENERALLY GOOD WHERE THE SOLUTION AND THE DOMAIN UNDERSTANDING HIGH
CONFERENCES & SYMPOSIA			GOOD ICEBREAKER BUT NOT EFFECTIVE BY ITSELF
INDIVIDUAL CONSULTING			VERY EFFECTIVE WITH LIMITED AUDIENCE
PATENTS & LICENCES			EFFECTIVE, RELIES ON PULL
COOPERATIVE RESEARCH			VERY EFFECTIVE WITH LIMITED AUDIENCE
SHOWROOM/TEST DRIVE			SOAKS UP A LOT OF DEVELOPMENT RESOURCES
CONSUMER REPORTS/TEST DRIVE			PROVIDES ONE STOP SHOPPING CAN BE DIFFICULT TO TURN LOOSE OF THE BABY
SMALL BUSINESS INITIATIVE	HIGH	HIGH	NEW FUNDING SOURCE, NEW PROCESSES, NOT DEVELOPMENT

JPL

TECHNOLOGY TRANSFER PROCESS



JPL

OPTIMAL SOLUTION
(MY VIEW)

- 3 REPLICATED DEMONSTRATION CENTERS
  - GSFC
  - JPL
  - LASP
- COORDINATED TESTING FOR MATURITY AND USABILITY OF TECHNOLOGY
- ONE CENTER (JPL) TO ADDRESS COMMERCIALIZATION ISSUES