UNIVERSITY OF CALIFORNIA, SANTA BARBARA

BERKELEY · DAVIS · IRVINE · LOS ANGELES · RIVERSIDE · SAN DIEGO · SAN FRANCISCO   SANTA BARBARA · SANTA CRUZ

SANTA BARBARA, CALIFORNIA   93106

REMOTE SENSING INFORMATION SCIENCES RESEARCH GROUP

SANTA BARBARA INFORMATION SCIENCES RESEARCH GROUP

PROGRESS REPORT AND PROPOSAL   - YEAR 4

TN-43

JANUARY 1, 1987

140 P.

PRINCIPAL INVESTIGATOR

DR. JOHN E. ESTES

CO-INVESTIGATOR / CONTRIBUTORS

DR. TERENCE SMITH
DR. JEFFREY L. STAR

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
OFFICE OF UNIVERSITY AFFAIRS
WASHINGTON D C   20546

GRANT NO. NASA NAGW - 455

ISRG

Table of Contents

Remote Sensing Information Sciences Research Group

Table of Contents

I. **Introduction**

John E. Estes and Jeffrey L. Star

II. **Reports**

Research Activities for the Information Systems Office
John E. Estes and Jeffrey L. Star

A Knowledge-based Geographic Information System
T. Smith, D. Peuquet, S. Menon, P. Agarwal

Requirements and Principles for the Implementation and
Construction of large-scale Geographic Information Systems
T. Smith, S. Menon, J. Star, J. Estes

Draft Preliminary Comparisons of the Pilot Data Systems
John E. Estes, David M. Stoms

Band Ratio Procedures for Maximizing Information Content
and Interpretability
R. Crippen, E. Hajic, J. Estes

Glossary of Terms and Abbreviations
J. Star, J. Estes, K. Lombard

Agricultural Investigations with an Information Systems
Context
J. Estes, J. Star, M. Caswell, D. Grice, C. Schmullius,
K. Bahk, D. Tong, M. Cosentino

University Research Libraries Group Geographic Information
System for Scholars
Jeffrey L. Star

III. **Budget**

IV. **Summary**

John E. Estes and Jeffrey L. Star

V. **Appendices**

Preliminary Design of a Farm Monitoring Geographic
Information System

Support for International Science

Browse in the EOS Era

Reports

Introduction

John E. Estes and Jeffrey L. Star

Introduction

The Office of University Affairs of the National Aeronautics and Space Administration (NASA) has signed a grant establishing a Remote Sensing Information Sciences Research Group (ISRG) at the University of California, Santa Barbara (UCSB). This document represents a progress report of work conducted under this grant (Grant # NASA NAGW-455) during the period May 1, 1986 to January, 1987, and proposal for the year starting May 1, 1987.

ISRG research continues to focus on improving the type, quantity, and quality of information which can be derived from remotely sensed data. As we move into the coming year of our research, we will continue to focus on information science research issues. In particular, we will focus on the needs of the remote sensing research and application community which will be served by the Earth Observing System (EOS) and Space Station, including associated polar and co-orbiting platforms. Research conducted under this grant has been used to extend and expand existing remote sensing research activities at UCSB in the areas of georeferenced information systems, machine assisted information extraction from image data, artificial intelligence and vegetation analysis and modeling.

As world population increases, there is an ever expanding need for systems and techniques capable of acquiring, integrating, and analyzing information concerning the extent, use of, and changes in the major components of the earth's surface. NASA is playing an important role in the development of systems

Introduction

such as EOS which have significant data acquisition capabilities. To achieve the full potential of such systems, however, requires that farsighted fundamental research be directed towards the scientific application of technologies upon which assessments may be made of both the current and changing status of the components of the biosphere, hydrosphere, lithosphere, and atmosphere.

The program of research, documented in this progress report, is being carried forward by personnel of the University of California, Santa Barbara. This report documents our accomplishments in what we consider to be a multiyear effort to prepare to take full advantage of the system's capabilities of the platforms and systems associated with Space Station (e.g., EOS). Through this work, we have targeted fundamental research aimed at improving our basic understanding of the role of information systems technologies and artificial intelligence techniques in the integration, manipulation and analysis of remotely sensed data for global scale studies. This coordinated research program is possible at UCSB due to a unique combination of researchers with experience in all these areas.

Efforts during the early years of this grant focused on the integration of existing research activities at UCSB and the initiation and conduct of a number of research activities with a variety of NASA centers. We have also worked on background assessments of research and technology, as well as beginning steps towards implementation of a Pilot Land Data System (PLDS) for NASA Headquarters. We continue to be involved in PLDS

Introduction

development efforts, largely through the Science Steering Group.
In addition, UCSB personnel have been involved with: the EOS Data
Systems Panel; Space Station Data User Working Group; Space
Station Operations Task Force; Science and Applications
Information Systems Working Group; Global Resources Information
Systems; the United Nations Environment Programs Global Resources
Information Database program; and, the Committee on Data
Management and Computation (CODMAC) of the National Academy of
Science.

In addition, during this past year we have received funds
from NASA Code EI to supplement ISRG activities. These funds
were proposed in September of 1985 to cover a range of tasks. We
have also received a major grant from NASA Code EE/EI to study
problems associated with Browse in the EOS ERA and have submitted
a proposal to Code EE to assist in system acquisition integration
and operation of the United Nations Environment Programs (UNEP)
Global Resource Information Database (GRID), Nairobi, Kenya
Facility.

The material which follows details ongoing work directly
aided by this grant. Several of the projects used this funding
as a catalyst to aid other NASA offices in the research, in the
integration of remotely sensed and other data into an information
sciences framework. The following sections discuss the details of
the projects dealing with:

* Studies Related to ISRG/Information Science Office
  Research;

* Artificial Intelligence and Geographic Information

Introduction

Systems, KBGIS- II: A Knowledge-Based Geographic Information System;

* Requirements and Principals for the Implementation and Construction of Large Scale Geographic Information Systems;

* NASA Pilot Data System Comparison;

* BROWSE of remotely sensed data in the 1990's;

* Nairobi Proposal;

* Band Ratios and Band Selection for Multispectral Image Generation;

* A Glossary of Data System Terms;

* An Agricultural Inventory Data Base for the Region of the Veneto, Italy; and,

* University Research Libraries Group, Geographic Information System for Scholars.

These projects are discussed in some detail in the following sections (and in associated Appendices) in terms of both ongoing and proposed activities for the coming year. This Progress report concludes with a section outlining the budget for this coming year's activities. The appendicies which follow contain: material expanding upon work from various sections of the report; publications which have been funded (in whole or part) by this grant; committees memberships held by our staff with relevance to information sciences; and, symposia and professional society papers are presented, related to work on this grant.

Research Activities for the Information Systems Office

John E. Estes and Jeffrey L. Star

Research Activities for the Information Systems Office

John E. Estes and Jeffrey L. Star

In 1985, UCSB ISRG proposed studies to directly support the efforts of NASA's Information Systems Office, Code EI. Our proposal discussed a number discipline areas which continue to be relevant to Code EI, and we have been in close contact with Code EI staff to determine which of these task areas provide the greatest return to NASA and the Information Systems Office. In the following text, we describe the different task areas and the work accomplished in each area.

The UCSB Information Science Group suggests the following topics for special studies for Code EI:

1. The interaction between the Code EI data pilots and the proposed EOS data system, and impacts on the university earth science remote sensing community;

2. University workstation activities and requirements as they relate to NASA Center and NASA Land and Ocean Data Pilots;

3. University roles in the planning and development of the OSSA Science and Applications Information System;

4. The impact of the imposition of uniform NASA data standards on university research;

5. University roles in the development of a Global Resources Information System;

6. Artificial Intelligence and the Image Understanding Process; and,

7. A Workshop on University Information Systems Requirements.

A good deal of planning, research, and operational effort has occurred through the Pilot Land, Ocean and Climate Data Systems, and somewhat independently, on the Earth Observing Systems data system.

Panels (with in some cases overlapping memberships) have convened on a number of occasions.

However, there has been little examination of the interactions between pilots and the EOS data system, nor what such interactions may imply for the academic Earth Sciences Community involved in remote sensing research. Questions will involve hardware, software, and communications compatibility issues. Links to NASA Centers and to other national (plus state, and local) as well as international data bases will also be involved. Universities will be an important part of the user community for the EOS and Pilot systems. However, what are the key issues involved in facilitating the use of data from these systems by university scientists? What needs to be accomplished to insure that this important segment of the university community can participate to the fullest extent in the research opportunities afforded by these systems?

NASA, through Code EI, Information Systems Office (ISO) has conducted studies over the past decade which have lead to the establishment of a number of pilot data systems. These pilot data systems are oriented towards the science driven data/information needs of users in defined disciplines under the NASA research umbrella. To date, pilot systems have been undertaken in the planetary, climate, oceans and land sciences areas.

The goal of these systems is to provide a variety of types of support for NASA and NASA related researchers. This support includes the maintenance of directories and catalogues, the formatting and preprocessing of data and the facilitation of communications between discipline scientists conducting NASA research.

NASA is also planning for Earth Observing Satellite System (EOS).

Research Activities for the Information Systems Office

EOS is designed to aid scientists in addressing the multidisciplinary earth science problems. The system must provide relevant information for detailed local studies, as well as for investigations aimed at an improved understanding of global dynamics. Recognizing the importance of such a system to scientific understanding, the EOS Science and Mission Requirements Working Group recommended that NASA should develop the Earth Observing System as an information system mission.

University researchers will make up an important element of the user community for information from both the Data Pilots and EOS. Yet, in the studies conducted to date, little attention has been focused on either the specific needs or the generic problems involved in achieving optimum university participation in such systems. We propose to conduct such a study.

Specifically, we have examined the operation of current NASA pilot systems as they pertain to interactions with university researchers. A preliminary version of this material is presented in a following section of this report. The report discusses the similarities and differences in the various pilots. It also discusses how the pilots have been organized in support of their diverse user communities and the level of satisfaction on the part of both "users" and "operators" of the pilot systems.

Our work in this area is continuing. In the coming months and into next year we will broaden somewhat the scope of this activity to examine what features of these pilots support the EOS Data System Concept as outlined in the Reports of the EOS Data Panel. The study will also look into how these existing pilots might be more effectively linked to the

broader NASA sponsored University Community.

Research Activities for the Information Systems Office

Science and Applications Information System

Another area of increased interest is large scale information systems. The Science and Applications Information System planning effort is currently being conducted within NASA Code E. This system, as currently envisioned, is being planned to answer the science and applications information systems requirements of NASA and NASA associated researchers working under Code E funding. This wide ranging effort is critical for effective and efficient use of EOS and other advanced systems being developed for use in the Space Station time frame of the 1990's and beyond.

We proposed to provide a university perspective to the SAIS planning process. We feel this is an important area where study and interaction with the University community is required. This has included working closely with NASA Headquarters and center personnel who are actively involved in the planning process. Questions regarding links to universities, transfer of data, data archiving, value added services, and standards are all of critical concern. The role of existing pilots and university beta test sites also should be addressed. For example, should SAIS planning be scenario driven? Expandability, heterogenous networks, and system evolution are all important concepts to the SAIS. As a result of our efforts in this area, Dr. Estes and Dr. Star are beginning to work closely on SAIS activities with NASA and other university personnel. Indeed, a major SAIS working meeting will be held the week of February 9, 1987 at UCSB. At this meeting, a cross-section of NASA and NASA-sponsored scientists will meet to be brought up-to-date on the Space Station program, and to focus on refining the structure of an SAIS system.

We at the ISRG will continue to support SAIS activities through reviews of project plans and attendance at SAIS science and working group meetings. In addition, our activities in the development of an image processing test bed will directly support our work, testing the SAIS concept for Telescience use of Space Station polar platforms in the EOS era.

Data Standards

Data standards have been discussed by many NASA and other federal agencies as well as science community committees for a considerable length of time. The imposition of standards implies that some change in operational procedures will be required to insure compatability. What impact will changes in data structures or headers have on university research and operations? To some extent, this can be examined within the context of changes which have occurred in Landsat data formats. The problem here, however, is much broader and the issues and impacts upon the university community need to be addressed.

We have been working with Jet Propulsion Laboratory personnel on data standards for the university earth science community. In this study, we are trying to determine both the advantages and potential disadvantages of establishing standards. The orientation of the study will be to consider a minimum set of standards which will improve and not impede universities' abilities to employ NASA acquired remote sensor information in an effective and efficient manner.

The study has examined one proposed data format family, based on a data description language. We have written software for both the IBM PC/AT and the VAX under Unix, to test the efficiency as well as the

logical structure of the proposed data standard. This part of the effort
will continue at a lower level in the second half of the current contract
year, as the standard and our software undergo review. Current practices
and procedures for acquisition, processing, analysis, reporting, and
storing of information will be reviewed. Areas where standards can be
employed to improve this process will be discussed, including
communication between sites and the development of geographic information
systems. We will also discuss the role of standardized software
interfaces. Emphasis will be placed on determining a minimum set of
standard formatting and processing procedures which will provide the
greatest benefit to the university community.

Global Resources Information System

In recent years, NASA has been moving towards conducting global
scale studies with time durations on the order of decades. Study
documents in the form of program plans have been developed for land-
related global habitability studies and for global biology studies. In
addition, the Space Station and its associated sensor platforms are being
justified in part on the basis of long term research studies on topics of
global significance and concern. The Earth Observing System (EOS)
Science Steering Group document clearly points out the need for such
studies. This same document also points out the need to consider EOS as
an information system.

ISO, at the urging of congress, has been developing the concept of a
Global Resources Information System (GRIS) for the past several years.
These studies have been lead by researchers at the Jet Propulsion
Laboratories. There is clearly a need for such systems. That need

extends from the use of such data as might be contained in such a system
in a policy making context, in a commercial context, and in a research
and development context. While the two former uses (policy and
commercial) are important, a case can be made that the research and
development aspects of the data in such a system would be extremely
significant. The question asked then is, who will be the real
beneficiaries of such a system? It is clear that among the top
beneficiaries will be university researchers. As one of these, we
propose a study of how universities can contribute to the
Research Activities for the Information Systems Office development of a
Global Resources Information System.

A strong case can be made that some of the greatest beneficiaries of
a Global Resources Information System (GRIS) will be found in the
university research community. Policy makers and federal agency
personnel will make extensive use of these data, but it will be
university researchers, doing advanced work in analysis and modelling,
who stand to reap significant near-term benefits from the global harvest
of data which can be provided by such a system. As such, it is important
to understand the roles which can be played by universities in the
development of a GRIS-type system and to address the issue of university
participation in the development and operation of Global Scientific
Information Systems.

As such, UCSB personnel have visited and will continue to work
closely with personnel of the UNEP/GRID program in Geneva, Switzerland,
as well as with personnel of the World Bank and NASA on Support for
Global Science and Applications research. An outgrowth of this work was
the placing of an intern at the World Bank this past year. Our student

worked toward his Master's Degree, and was responsible for operating an
image processing system at the World Bank Headquarters in Washington,
D.C. Our student, Mr. Paul Lefebvre, participated in training of Bank
employees as well as participating directly in the development of
materials for Bank projects in developing nations. This was part of an
effort at the World Bank, evaluating the utility of remotely sensed data,
as well as developing some in-house experience with the information
systems requirements. At the present time, we are in the process of
negotiating another intern for the World Bank. Bank personnel have
expressed considerable interest in this. We hope our efforts in this
area are successful.

In addition, we at UCSB working with personnel from Code EE and the
National Science and Technology Laboratory at Bay St. Louis, Mississippi,
have proposed an effort to support the upgrading of computed facilities
and the UNEP/GRID facility in Nairobi, Kenya. The proposal for this
effort is contained in an appendix to this report. We remain interested
in and are actively pursuing improved insights into the institutions,
techniques and methods currently available for the conduct of large scale
and global science. We feel our activities in this area are progressing
well. This work will continue into the coming year through our proposal
efforts. In addition, we will continue to examine ways to improve .
science and applications data linkages on the international level. This
work is important if we are to move toward the goals laid out by the
Earth Science Steering Committee in the Earth System Science Report.

Artificial Intelligence

Image analysis is a process which is integral to all remote sensing

10

systems. Despite advances in computer-assisted image processing, the information extraction potential inherent in digital image processing lags behind the level of information extraction exhibited by manual image interpretation techniques. An examination of the processes involved in computer-assisted versus manual image interpretation reveals that computer assisted techniques rely largely on a single element, specifically tone or color, and a primitive form of texture; whereas manual interpretation techniques incorporate much higher level inputs (e.g., context).

Since concepts from the field of artificial intelligence (AI) have proven useful in applications similar to those found in the interpretation and analysis of remotely sensed images, UCSB has been conducting research for several years to identify specific research directions in which AI techniques can be applied to remote sensing with potential high returns. Much of this work has been driven by Prof. T. Smith, co-investigator on this grant. This leads us to propose a study in which some of the fundamental research required in the development of a computer assisted interpretation system is conducted.

* Since the data inputs to an image interpretation system are by necessity varied and complex, a reasoning model must first be developed that incorporates the various data forms and models their interactions and preferred order of occurrence. This involves the assemblance and structuring of the knowledge of expert image interpreters, and the formation of information structures which allow efficient storage and access to the varied data forms. The data may be digital (image or other collateral data) or semantic in nature, and may concern either the applications area or background information concerning the region to be analysed.

* In order to optimize the efficient search through a large and complex knowledge base, a system must have some way of prioritizing the processing instructions. One way to develop this prioritization is by conductiong a case study of the computer assisted processing and analysis of land use/land cover classification studies, in order

to develop reasonable a priori probabilities for the initial program. These a priori probabilities could then be incorporated into the decision model to indicate the preferred processing flow.

* Lisp is often cited as the language of choice in an AI program. However, in a digital image processing context, where simply an excess of input/output initiations can produce an increase in the cost of running an image processing program by an order of magnitude, Lisp may not be the most appropriate programming language. A study is needed that will examine the unique problems involved in the data structuring, data interaction and digital processing to determine which programming language would be optimal for that data or instructional module. The languages which would be examined for this study would include Lisp, Prolog, C and Fortran. It is anticipated that different modules would be optimally represented in different languages. This could be followed by a cost/benefit analysis to integrate the development and operational considerations involved in a computer-assisted analysis system in order to reveal the programming direction to task in the design of such a system.

Conclusion

In conclusion, we believe that our work in this area represents a particularly efficient mechanism for the UCSB ISRG to support NASA's Code EI. With the decision on which of the technical and scientific tasks are most important remaining with Code EI staff, we are able to respond to changing needs and requirements at Headquarters on a timely and efficient basis. We look forward to this relationship, and to the continuing opportunity of supplying a university perspective to Code EI.

Draft Preliminary
Comparison of Pilot Data Systems

John E. Estes and David M. Stoms

Dratt Preliminary Comparison of Pilot Data Systems

NASA has funded pilot data systems in the ocean, climate, planetary and land sciences disciplines. The data pilots are designed to examine the acquisition, storage, analysis, and use of remotely sensed data in the different disciplines. To date, there has been little effort spent to examine the important features of each of these test efforts, to compare and contrast the different approaches taken, and to consider the commonalities and ditferences between them.

The tollowing chapter is a preliminary attempt at just such a comparison. As part of our efforts in 1987, we propose to retine this document, by examining the successes of each of the pilot programs, and identifying the areas where additional work could be beneficial to NASA's new efforts in the information sciences.

It is our feeling that such work is critical it we are to use the existing pilots as the foundation, the building blocks, upon which the EOS data system is to be developed, as some have suggested.

Draft Preliminary
Comparison of Pilot Data Systems

prepared for NASA Headquarters

Principal Investigator

Dr. John E. Estes

Co-investigator

David Stoms

University of California, Santa Barbara
Remote Sensing Research Unit

Draft
January, 1987

Draft Preliminary

COMPARISON OF PILOT DATA SYSTEMS

1. Background

The National Aeronautics and Space Administration (NASA) has developed four information systems to serve as prototypes for demonstrating principles, techniques, systems, and capabilities for the assembling, storing, management, and dissemination of discipline-specific remotely sensed data sets. One of these pilot systems, the Pilot Ocean Data System (PODS) has been implemented by NASA as NODS, the NASA Ocean Data System. The Pilot Climate Data System (PCDS), while also considered operational, still carries the term 'pilot' in its title. Different codes within NASA have lead responsibility for each system. To date, a thorough comparison of important features of the four data systems has not been made. This paper addresses that situation by comparing system components. Hopefully, the systems will benefit from this comparison by drawing on each others' experience.

Each data system supports major discipline areas within NASA--lands, oceans, climate, and planets. As stated above, NODS and PCDS are currently operational. The Pilot Land Data System (PLDS) and the Planetary Data System (PDS) will have preliminary versions operating in 1987. The objective of these pilot projects is to make the large (but widely distributed) collections of satellite and ancillary observations of interest to NASA and NASA-related discipline scientists more readily accessible to researchers who may not even know what data are available or how to get hold of them. Future data systems such as the Science

and Applications Information System (SAIS) and the Earth Observations System (EOS) will integrate the pilot systems and expand upon them.

The need for improved management of satellite and in-situ data sets has emerged from many related factors. The public has made enormous investments in space-related research. Large volumes of data accumulated represent a national treasure that cannot be replaced and should be preserved. Concurrently, the potential scientific value of the data means these data sets should be accessible to researchers. Typically, however, once the active phase of the research satellite mission has ended, the resources available to manage the data diminishes. Data is then archived in scattered locations, in non-uniform formats, with undocumented processing, and therefore is not adequate for use. It was in recognition of this potential loss of priceless data, combined with the anticipation of more data from the Earth Observing System on Space Station in the 1990's, that induced NASA to promote these four pilot systems.

The discussion begins with a brief description of each system--it's mission, research user community, data sets, and test bed applications. Following that, the basic functional components of each system will be compared. This task is complicated by differences in terminology between systems and by differences in their hierarchical organizations, and by differences in the details in the documentation. Next, the individual functions will be covered in somewhat greater detail (e.g., means of access, user interface, data manipulation and display, and other functions such as bibliographic search and on-line help).

The primary intent of the material presented here is to gain a greater appreciation for the diversity of approaches in establishing the NASA data pilots and to gain an improved understanding of the fundamental differences and basic similarities in these systems. Further work will be required to address questions such as which parts of which systems are particularly well suited for a given task or function. Can standard approaches to problems evolve out of these systems; how can lessons learned from one system be incorporated effectively into others; what parts of these systems can be used to form an EOS data system; is that even

appropriate; and so on?

2. Pilot Data System Descriptions

Pilot Climate Data System

The collection of climate research data has been a major focus within NASA. PCDS was the first NASA data system to provide on-line capability to users. It was designed and scoped in 1980 and began test operations in June, 1982. The newest revision, Version 3.5, was installed in September, 1986. Current research applications being supported include Statistical Climatology, International Satellite Cloud Climatology Project, Global Ozone Distribution, Earth Radiation Budget Studies, Land-Surface Climatology, and Solar Flux. By the end of 1987, it is expected that PCDS will be declared operational, with a name change to the NASA Climate Data System (NCDS).

The goal of PCDS is to develop a system for locating, obtaining, manipulating, and displaying climate data of interest to NASA's research community. Specific objectives are to develop a technically advanced system, to obtain and incorporate the most important data sets, and to support a broad range of users.

PCDS operates on a VAX 11/780 at NASA's Goddard Space Flight Center (GSFC), housing both the software and NASA and non-NASA data sets. The system can be accessed 24 hours per day, seven days a week, except during scheduled maintenance. While PCDS has no specific science advisory panel participating in its management, the scientific community has interacted frequently, including two workshops held at GSFC. Mary G. Reph is the PCDS Project Manager. The contact address is:

Pilot Climate Data System
NASA Goddard Space Flight Center
National Space Science Data Center
Data Management Systems Facility
Code 634
Greenbelt, MD 20771

University of California, Santa Barbara

(301) 286-5037

NASA Ocean Data System

The data system to serve the oceanographic community began as the Pilot Ocean Data System (PODS) but was declared operational as NODS. The latest major system release was installed in February, 1986.

NODS was designed to give the oceanographic community access to a variety of remotely sensed and in situ data and functions to preview the data.

The system is managed at the Jet Propulsion Laboratory (JPL), currently providing remotely sensed data sets, both on-line and off-line and also on-line in-situ data sets. NODS is available 24 hours a day, seven days a week, except during scheduled maintenance. Current recommendations from the Advisory Panel of the Satellite Ocean Data Systems Science Working Group emphasizes that NODS should be part of a distributed data retrieval system and not focus on data analysis. Panel members felt that analytical functions are already available and that work on incorporating them into NODS would detract from the primary effort of data retrieval. The Project Manager is Dr. Eni Njoku and the Project Scientist is Dr. Victor Zlotnicki. The contact address is:

> The NASA Ocean Data System
> Jet Propulsion Laboratory
> MS 202-101
> 4800 Oak Grove Drive
> Pasadena, CA 91009
> (818) 354-5036
> FTS 792-5036

Planetary Data System

PDS originally began as two separate tasks in 1983. The Pilot Planetary Data System was established to evaluate information systems and planetary science research methodologies. The other task, called the Planetary Data System, began defining the requirements of

an operational planetary data system. The two projects were later merged and named the Planetary Data System. Version 1 of the system is scheduled to begin operating in mid-FY 87, and Version 2 by the end of FY 89.

The overall goal of PDS is to serve as curator of planetary data, providing both increased longevity of data and knowledge of their meaning, and to improve the organization of and access to these data in order to facilitate scientific investigation and analysis. A more specific objective of PDS is to develop and implement a system, based on the recommendations of the Committee on Data Management and Computation (CODMAC), which can provide easy access to planetary data and to information about the data to the scientific community.

Project responsibility is assigned to JPL, where the high-level catalog and archives are maintained on a VAX 11/780. It is designed to be a distributed system with a central node at JPL which maintains the high level catalog and the archives on a VAX 11/780. Discipline nodes, located at installations actively doing planetary research, maintain mature data sets and expert disciplinary knowledge. The discipline nodes are organized around four subdisciplines—geology, rings, atmosphere, and fields and particles. Mission nodes provide the interface between Planetary Flight Centers and PDS. A Scientific Working Group participates in the oversight of PDS. Dr. William Kurth, a research scientist at the University of Iowa, has replaced Dr. Ray Arvidson as Project Scientist. Project Manager for PDS is Dr. Tom Renfrow at JPL. Contact address is:

> Planetary Data System
> Jet Propulsion Laboratory
> 4800 Oak Grove Drive
> Pasadena, CA 91009
> (818) 354-6347
> FTS 792-6347

Pilot Land Data System

University of California, Santa Barbara

Planning for a pilot data system to serve earth scientists began in 1983. Preliminary users requirements have been drafted. An initial operating version of PLDS, called Build 1, is scheduled to go on-line in July, 1987. Development will continue through 1989. The three test bed projects in the development phase are the International Satellite Land Surface Climatology Project Retroactive Analysis Project (ISLSCP-IRAP), the First ISLSCP Field Experiment (FIFE), and the Sedimentary Basins Project (SEDBAS). A fully operational Land Data System is anticipated to be on-line around 1991.

PLDS was created to provide a limited scale demonstration of an information system that would offer scientists the ability to archive, locate, transfer, integrate, and manipulate data in a distributed fashion.

PLDS is managed at NASA's GSFC, although it is not yet accessible to users. Data storage will be distributed at existing archives such as EROS Data Center, USGS/ESDD, USGS/WATSTOR, and NOAA/NEDRIS. In the preliminary demonstrations of PLDS, scientists can access these archives but will not be able to query or command satellite sensors directly. This type of capability is being examined for the EOS time frame, however. Dr. Paul Smith at GSFC is Project Manager; Dr. Robert Price, also at GSFC, is Project Scientist. The address is:

> Pilot Land Data System
> NASA Goddard Space Flight Center
> National Space Science Data Center
> Data Management Systems Facility
> Code 634
> Greenbelt, MD 20771
> (301) 286-5037

3. Functional Components

Organizing a comparison of pilot data systems is not a clear cut task. The pilots are in different stages of development, (see Figure 1), so the level of specificity varies.

Each system uses a slightly different terminology for similar functions. Some call a function a

Figure 1. Pilot Data System Development Timelines


```
        78      80      82      84      86      88      90      92

  PCDS  ▽-------O-------------------◇------------------

  NCDS  -------------------◇------------------------------

   PDS          ▽--------------------O-------◇----------

  PLDS          ▽-----------------O-----------------◇----

        ▽  Planning
        O  On-Line
        ◇  Operational
```
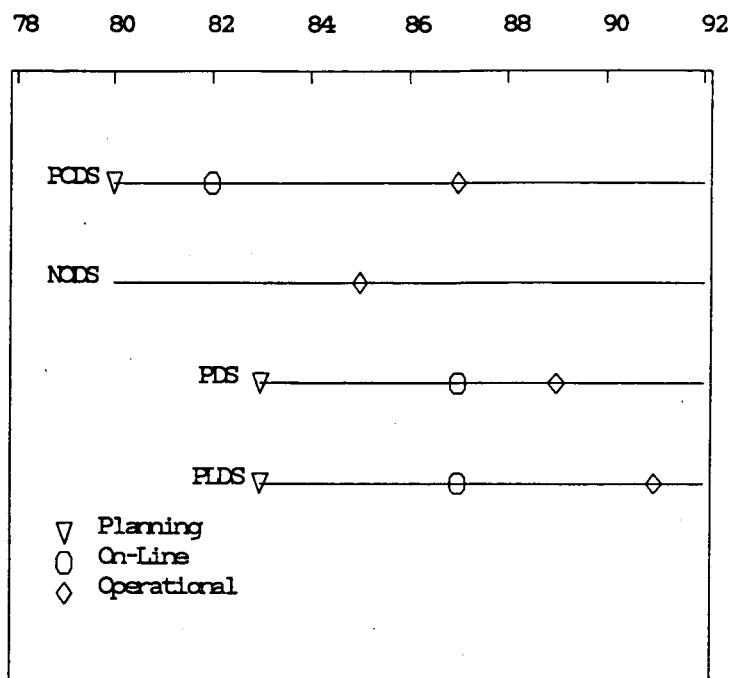
component or a major subsystem, while others simply consider the same function to be an element of a larger component. In other words, the functional hierarchies are quite different. Indeed, throughout this entire area there is a real need to establish, in so far as possible, a glossary containing the agreed definitions of a variety of terms. Such a glossary would improve and facilitate communications among data system scientists and technicians, discipline scientists, and researchers. To allow comparison, this paper focuses on and identifies six typical subsystems of a data system, plus an 'all other' category for miscellaneous and administrative functions. These subsystems or components are:

1. User Interface

2. Catalog and Directory

University of California, Santa Barbara

3. Data Validation

4. Data Retrieval

5. Data Manipulation

6. Data Display

7. Other

Each of these subsystems can in turn be broken down into subunits in a variety of ways. These will be discussed in more detail in the sections which follow. Table 1 summarizes these functional components of the four pilot data systems. Cases identified with an asterisk indicate that the system has some capability in the given function but that it is not considered a separate subsystem.

Note that NODS has no subsystem for data manipulation. As mentioned above, the Science Working Group overseeing NODS found that at present, adequate capability in oceanographic data processing and analysis and wanted to see NODS focused on data retrieval. Also notice that there is no data acquisition subsystem listed above. In all four data systems, the mission planning aspects are detached from the data management functions. The focus of the pilot studies is on management of data from past and upcoming missions, rather than on recommending missions to collect missing data.

The details of each of these subsystems for each data system will be compared in the following sections. These comparisons should be taken with some caution, however. These are pilot systems, the details of which are still evolving. The absence of a particular capability or function indicates only means that, at this time, it was not found in the available documentation. It cannot be asserted with certainty that the function has been consciously or unconsciously omitted by system planners.

User Interface

University of California, Santa Barbara

Table 1. Functional Components/Subsystems				
Function	PLDS	NODS	PCDS	PDS
User Interface	*	Menu/ Command Interface	*	Access Data
Catalog/ Directory	Catalog, Directory	Global On-Line Data Catalog	Catalog, Inventory	Inspect Data
Data Input/ Validation		*	*	Prepare Data
Data Retrieval	Data Reduction	Data Archive	Data Access	Inspect Data
Data Manipulation	Data Reduction		Data Manipu- lation	Inspect Data
Data Display	Graphics	Data Display & Product Generation	Graphics	Inspect Data
Other	*	*	*	*

* system has this capability but it is not as yet nor planned to be a major component

PLDS = Pilot Land Data System
NODS = NASA Ocean Data System
PCDS = Pilot Climate Data System
PDS = Planetary Data System

The User Interface component generally refers to how a scientist can access the data system and interact with it. We will consider three aspects. First, the actual user interface-- how the user tells the system what to do--is described. Then the variations in the means of accessing the system and the user hardware requirements are compared.

As shown in Table 2, all four pilot systems allow both menu-assisted operations for beginners and a command language for experienced users. PLDS further proposes allowing natural language queries in some future revision. Most offer an on-line Help function, and

PCDS even has an on-line tutorial session. All four pilot systems allow queries by some combination of parameters such as date, location, data type, platform, sensor, and so on. However, the details are not given to the same degree of specificity, so an adequate comparison cannot be made at this time.

Table 2. User Interface Capabilities				
Capability	PLDS	NODS	PCDS	PDS
Menu	•	•	•	•
Command Language	•	•	•	•
Natural Language	?			
On-Line Help		•	•	•

PLDS = Pilot Land Data System
NODS = NASA Ocean Data System
PCDS = Pilot Climate Data System
PDS = Planetary Data System

The data systems vary in their means of access (see Table 3). All four will allow use of a packet-switching network like TELENET. All but NODS allow direct dial-up over a modem, and all but PLDS use the Space Physics Analysis Network (SPAN). PLDS and PDS will include a local area network, and PLDS anticipates future use of satellite communications.

Hardware requirements are intentionally kept simple (see Table 4). Most systems only require the user to have any ASCII terminal or a personal computer with emulator/communications software and a modem. A scientific workstation is the expected device for PLDS. Printers, plotters, or graphics displays are optional, although, of course, the capability to view output displays depends on some combination of them.

Catalog and Directory

University of California, Santa Barbara

Table 3. Methods of Access				
Method	PLDS	NODS	PCDS	PDS
Local Area Network at Central Node	●			●
Local Area Network at Satellite Node				●
SPAN Network		●	●	●
TELENET Type Network (packet switching)	●	●	●	●
Direct Dial/Modem	●		●	●
Satellite Communications	?			

PLDS = Pilot Land Data System
NODS = NASA Ocean Data System
PCDS = Pilot Climate Data System
PDS = Planetary Data System

Table 4. Required User Hardware				
Device Type	PLDS	NODS	PCDS	PDS
Terminal or PC		●	●	
Modem		●	●	
Scientific Workstation	●			
Printer		*	*	
Plotter		*	*	
Graphics Display		*	*	

● = required hardware
* = optional hardware

PLDS = Pilot Land Data System
NODS = NASA Ocean Data System
PCDS = Pilot Climate Data System
PDS = Planetary Data System

University of California, Santa Barbara

The catalog or directory component clearly illustrates the lack of uniformity in pilot data systems. PLDS tentatively has separate Directory and Catalog subsystems. The Directory will contain detailed information about the data sets, while the Catalog will have more general descriptions about the data collections. The Global On-Line Data Catalog in NODS has two levels. The top level, comparable to PLDS' Directory, can be queried on data parameters such as platform, sensor, location, and time. The lower level includes more general information about the data sets, similar to PLDS' Catalog subsystem. PCDS has two separate subsystems like PLDS does, but calls them Inventory and Catalog. In PDS, the Inspect Data subsystem includes a catalog function as well as the data manipulation and display functions.

Data Validation

Only PDS has a specific component for inputting and validating data. Quality control checks, specified by the planetary science community, are performed. Errors are corrected internally, when possible, or else the data are returned to the originator. Of course, the data system user will not interact directly with this subsystem. However, it will be crucial to the long-term success of future information systems that data sets are carefully chosen and that their quality is validated and described in the catalog.

Data Retrieval

Retrieving data from the archives is a significant aspect of each pilot system. Each has a subsystem dedicated to data retrieval, except for PDS which includes it in a comprehensive Inspect Data component. Basically, all systems allow queries by mission or platform, by sensor, by date and time, by geographic location, or by data type, or some combination of logical 'and' and 'or' operators.

University of California, Santa Barbara

Data Manipulation

The four pilot systems vary significantly in their emphasis on analytical functions (see Table 5). In PLDS, it is anticipated that functions for radiometric and geometric corrections and for image subsetting will be provided. Other functions such as statistics, modeling, image processing, and geographic overlay may also be available. NODS, on the other hand, will provide almost no analytical capabilities, but will focus instead on data base management aspects. PCDS is limited to customizing data sets for further processing at the user's node or preparing data for the graphics subsystem using statistical functions. PDS will incorporate functions for statistical and mathematical analysis, sampling, and elementary image processing. The responsibility for software development in PDS is assigned to the discipline nodes.

Table 5. Data Manipulation Capabilities				
Capability	PLDS	NODS	PCDS	PDS
Statistics	●		●	●
Data Sampling				●
Math Functions			●	●
Modeling	●			
Image Processing	●			●

PLDS = Pilot Land Data System
NODS = NASA Ocean Data System
PCDS = Pilot Climate Data System
PDS = Planetary Data System

Data Display

The final output of the data systems is their real objective. All four systems identify this as an essential aspect by making a separate subsystem (although PDS includes other elements within its Inspect Data subsystem). As shown in Table 6, the displays include tabular and textual data, graphics, imagery, and in NODS, bibliographic abstracts. Naturally, the

University of California, Santa Barbara

capabilities are dependent on the user's hardware, although custom products or tapes can be produced at the central node and shipped to the user. Some graphic products may necessarily be performed at the central node because of the transmission times. For instance, it would take 45 minutes to transmit a 512 x 512 pixel image with 8-bit pixels over a 1200 baud line (Jet Propulsion Laboratory, 1986).

Table 6. Display/Graphics Capabilities				
Capability	PLDS	NODS	PCDS	PDS
Text	●	●	●	●
Abstracts		●		
2D Graphics	●	●	●	●
3D Graphics			●	●
Imagery	●	●		●
Browse	●	●	●	●

PLDS = Pilot Land Data System
NODS = NASA Ocean Data System
PCDS = Pilot Climate Data System
PDS = Planetary Data System

Browsing capability is proposed for all four data systems. NODS will have Browse functions both for looking at examples in the Catalog subsystem and at data in the Archive subsystem. In PCDS, Browse refers to scanning detailed catalog descriptions. PDS will utilize optical videodisks for electronic Browse of planetary images stored at regional facilities.

The role of a Browse function is unresolved and may well be different for each data system because of the type of data and because of differences in user requirements. The tradeoff between storage of large numbers of preprocessed images and the time time of processing only on request need to be examined. Related to this issue is determining what types of browsable imagery the scientific community wants, realizing there are a vast number of possibilities. These issues are being addressed in a current research study at the Remote Sensing Research Unit of the Department of Geography, University of California, Santa Barbara.

Other Functions

This section describes the capabilities of the pilot data systems in several miscellaneous areas, including bibliographies, electronic mail, and administrative details (see Table 7).

Table 7. Other Functions				
Function	PLDS	NODS	PCDS	PDS
Bibliography	*	Biblio-graphy		*
Electronic Mail	*			*
Adminis-tration	*	*	*	Order Data, Distribute Data, Administer Data

* system has this capability but it is not as yet or planned to be a major component

PLDS = Pilot Land Data System
NODS = NASA Ocean Data System
PCDS = Pilot Climate Data System
PDS = Planetary Data System

PLDS and PDS will have some bibliographic data available. Only NODS, however, actually has a separate subsystem just for bibliographic data. The bibliography can be queried by author, title, subject, and so on, with the results of the query sent to the user.

Both PLDS and PDS anticipate some form of electronic mail to allow communications between users and between the system and the users. For instance, news of system revisions and new data sets can be distributed to all users.

Administrative functions such as ordering and distributing data will be provided by all four systems. PDS emphasizes this aspect with three of its five subsystems directed towards these administrative details.

University of California, Santa Barbara

4. Conclusions

The four pilot data systems in NASA have been developed more or less in parallel and somewhat independently. They are being implemented by different NASA centers and overseen by separate scientific steering committees. Hopefully, this report has provided a greater appreciation for the diversity of approaches in developing the NASA data pilots and for the differences and similarities in the products.

A major task remaining prior to integration of the pilots into an Earth Observations Information System is to evaluate each of pilot system. Important questions remain as to whether researchers will actually use and benefit from distributed data systems. The development of a Browse capability, currently underway at the University of California, Santa Barbara, will be crucial for allowing scientists to preview data at their local terminals to determine if it meets their needs. To accomplish this task will first require a survey to find out how researchers decide when data is adequate, e.g. by visually inspecting raw or transformed imagery, by summary statistics, etc.

Nevertheless, NASA has already made great strides towards designing a vehicle to make satellite data accessible by researchers in the land, ocean, climate, and planetary sciences. The wealth of data from past and future missions may begin paying dividends far beyond the scope of the original research mission objectives, particularly when data from different sources can be integrated synergistically.

5. Bibliography

Estes, John E., Jeffrey L. Star, Phillip J. Cressy, and Michael Devirian, "Pilot Land Data System," *Photogrammetric Engineering and Remote Sensing*, vol. 51, no. 6, pp. 703-709, 1985.

Jet Propulsion Laboratory, *Planetary Data System: Version 1.0, System Specification*, p. 234, July 16, 1986.

Klose, J.C., *NASA Ocean Data System: Presentation to CODMAC*, p. 19, January 8, 1986.

NASA Goddard Space Flight Center, *Pilot Climate Data System*, p. 14, date unknown.

NASA Goddard Space Flight Center, *Pilot Climate Data Base Management System*, p. 4, May, 1983.

NASA Goddard Space Flight Center, *Proceedings of the Second Pilot Climate Data System Workshop*, January 29-30, 1986.

NASA Office of Space Science and Applications, *Information Systems Newsletter*, vol. 7, p. 39, September, 1986.

Renfrow, Tom, *Planetary Data System: CODMAC Review*, p. 11, January 8, 1986.

Science Applications International Corporation, *Earth and Environmental Science in the 1980's: Environmental Data Systems, Supercomputer Facilities, and Networks*, October, 1986. Contract NASW-3622

Band Ratio Procedures for Maximizing Information Content and Interpretability

R. Crippen, E. Hajic, J. Estes

Band Ratioing Procedures for Maximizing

Information Content and Interpretability

This is a draft of a paper, being prepared for submission to
a refereed journal. The subject of this paper is the use of band
ratios to improve the interpretability of remotely sensed data.
In particular, with new aircraft sensors such as AIS and AVHRIS,
and proposed spaceborne sensors such as HIRIS, future image
analysts will have many more spectral channels than in the past,
and thus, many more possible band ratios. This paper examines
the desireable characteristics of band ratios, and suggests
algorithms for their production.

BAND RATIOING PROCEDURES

FOR MAXIMIZING INFORMATION CONTENT

AND INTERPRETABILITY

by

Robert E. Crippen

Earl J. Hajic

John E. Estes

Remote Sensing Research Unit

University of California

Santa,Barbara, CA 93106

PRELIMINARY DRAFT

INTRODUCTION

Band ratioing is an image enhancement procedure in which the radiance values of two bands of a multispectral image are divided, one by the other on a pixel by pixel basis, in order to produce a pseudoimage known as a "ratio image". Descriptions of specific band-ratioing procedures are scarce in the literature, and it is clear that many researchers have produced ratio images that fall short of maximizing overall information content and interpretability. This paper reviews the logical basis for band ratioing, discusses the desirable characteristics of ratio images, and compares several effective procedures for the production of them.

The use of ratio images has been widespread and highly touted in the analysis of remote sensing imagery data, especially for lithologic studies (e.g., Vincent, 1972; Rowan et al., 1974,1977,1982; Goetz et al., 1975; Raines et al., 1978; Podwysocki et al., 1983) and vegetative studies (e.g., Maxwell, 1976; Tucker, 1979; Jackson et al., 1983). Band ratios enhance image information regarding the spectral reflectance properties of surface materials by subduing those factors of surface radiance that are correlated between bands, namely albedo and illumination intensity.

Useful ratio images are made from two bands in which the various features of interest differ non-proportionally in their spectral reflectivities. These bands can be

selected on theoretical grounds from knowledge of target spectra, or by evaluating test images of selected sites. The topics of band and ratio selection are specifically addressed by Crippen et al. (1987). The concern of this paper is instead the calculation and digital representation of the ratio image.

RATIONAL FOR BAND RATIOING

The logical basis of digital band ratioing was presented by Kreigler et al. (1969). It can be readily illustrated. On a bi-spectral plot (Figure 1) pixels of a homogeneous surface material in rugged terrain generally fall along a straight line. The slope of such a line approximately equals the ratio of the two band reflectivities for the surface material, and the position of a pixel along the line is a function of the solar illumination intensity for that pixel. If pixel values are calibrated so that zero in each band represents zero ground radiance (the line projects through the graph origin), then the ratio values of the pixels along the line are fairly uniform, equal to the line slope, and independent of illumination intensity. Since spacial variations in illumination intensity are largely controlled by topographic orientation relative to the sun position, band ratios are largely free of topographic effects. Conversely, since pixels of materials of differing reflectance properties commonly plot along lines of differing slopes, ratio values

are strongly related to surface material composition.
Figure 2 provides an example of two bands and their ratio
image.

Where topography is not prominent, band ratioing can
still be useful since it also reduces spatial variations in
signal strength attributable to heterogeneities in
atmospheric transmittance and to look-angle effects. Such
variations will be removed by band ratioing to the extent
that they are constantly proportional between bands across
the image.

The logical basis of ratioing can also be described in
terms of the various components of sensor-recorded radiance
measurements. Equation 1 (after Kowalik et al., 1983)
describes, to the first order, a radiometric measurement by
a nadir-viewing sensor. The measurement consists of four
components *:

  (a) ground reflectance of direct solar illumination
  (b) ground reflectance of skylight
  (c) atmospheric path radiance
  (d) sensor calibration offset

The first two components are modified by atmospheric
transmittance, and all components are adjusted by the sensor
gain factor. It is assumed that the ground is a lambertian
surface.

footnote----------------------

* If one defines pixels as rectangular areas having
  sharp boundaries, then a pixel "cross-radiance" component
  must also be included. If one instead accepts the
  indistinct, overlapping, <u>true</u> spatial character of pixels
  (as determined by both sensor and atmospheric optics), then
  this additional component is meaningless.

$$L = \frac{S\,T\,\rho\,Hd\,\cos\alpha}{\pi} + \frac{S\,T\,\rho\,Hs}{\pi} + S\,A + S\,O \qquad [1]$$

L = recorded radiance

S = sensor gain factor

T = atmospheric transmittance

ρ = surface reflectance

Hd = direct illumination intensity

$\cos\alpha$ = cosine of the incidence angle

Hs = skylight illumination intensity

A = atmospheric path radiance

O = sensor calibration offset

All variables are specific to each wavelength band except
the ground surface orientation term, $\cos\alpha$.

The underlying premise of band ratioing is that
recorded radiances can be adjusted so that their ratio is
approximately proportional to the ratio of surface
reflectances, ρ_a/ρ_b (for two bands "a" and "b"). Clearly,

the approximation is unlikely to be close if the atmospheric
path radiance or sensor calibration offset components are
significant. The approximation is good, however, if (1) the
direct illumination, atmospheric transmittance, and sensor
gain are essentially uniform for pixels within each band, or
are constantly proportional across the bands to be ratioed,
(2) skylight either provides a negligible contribution to
total illumination or has spectral proportions similar to
the direct illumination, and (3) the atmospheric path
radiance and sensor offset components can be estimated and
removed (discussed below). If these conditions hold then
all terms other than reflectance will reduce to a constant,
c, by which the ratio of recorded radiances, adjusted by
(3), are linearly related to the ratio of surface
reflectances (Equation 2).

$$\text{ratio of adjusted radiances} = \frac{L_a - (S_a A_a + S_a O_a)}{L_b - (S_b A_b + S_b O_b)} \approx c \frac{\rho_a}{\rho_b} \qquad [2]$$

for bands "a" and "b"

Since the cos α (solar incidence) terms are equivalent
between bands, they are eliminated by band ratioing (given
the above stated conditions). Thus, unlike the original
bands, properly derived ratio values are generally not
correlated with topography. Instead, ratio values are
almost entirely a function of surface reflectivities and,
therefore, surface composition.

RATIO IMAGE SIGNAL-TO-NOISE RATIOS

Many researchers have noted that noise commonly appears greater in ratio images than in the original bands. This is largely attributable to a proportionally greater reduction in image signal than image noise during the ratioing process. Image signal typically has a strong positive correlation (such as in Figure 1). Ratios therefore measure the bi-spectral data in a dimension that commonly has minimum signal variability. Noise, on the other hand, typically shows no interband correlation and measurements of it in the "ratio dimension" are approximately equal in magnitude to the average noise in the two original bands. Thus, ratioing reduces the signal but not the noise, and this results in a smaller signal-to-noise ratio. Additionally, some new quantization noise is added when the ratio values are represented by integers.

THE IDEAL GENERAL-PURPOSE RATIO IMAGE

In general, the following characteristics are desirable for ratio images:

(1) The input radiance values should be corrected for atmospheric and sensor-offset additive-radiance terms so that the ratio values are closely representative of the reflectance properties of surface materials.

(2) The distribution of ratio values should be positioned and scaled to fill the quantization range, with only minor

saturation at the range limits.

(3) The distribution of ratio values should tend toward histogram equalization (histogram flattening) so as to maximize, on average, the distinguishability of each pixel from all other pixels (maximize image entropy).

(4) The ratio image should be invertible, so that information content does not change when the numerator and denominator bands are exchanged. Being invertible means that a ratio image is the negative of the ratio image produced by inverting the ratio. Invertible ratios are, on average, symmetrical about a central (median or mean) value, and deviations from this symmetry are attributable to image features and not to the ratioing procedure.

In cases where only anomalously extreme values are of interest, some of these characteristics (such as histogram equalization) may not be desireable. However, for general purpose ratioing, the above characteristics tend to maximize information content and present it in a readily interpretable form.

RAW RATIOS: PROBLEMS

An inherently problematic yet still common procedure for producing ratio images is to divide the raw (unadjusted) pixel values of one band by those of another band to produce a "raw ratio" (which is then multiplied by a scaling factor). This procedure fails to compensate for the

additive terms in Equation 1. Since these terms are largely independent of the reflectance properties of the surface materials, the resultant ratio values are not necessarily closely related to surface composition. As expressed algebraically in Equations 3a and 3b, ratio values are not constant, despite proportionally equal surface radiances (Ra and Rb), if the additive terms (Aa and Ab) are not removed.

$$\frac{Ra}{Rb} : \frac{15}{10} = \frac{30}{20} = \frac{45}{30} \quad , \quad [3a]$$

but

$$\frac{Ra + Aa}{Rb + Ab} : \frac{15 + 8}{10 + 3} =/= \frac{30 + 8}{20 + 3} =/= \frac{45 + 8}{30 + 3} \quad . \quad [3b]$$

Commonly, the additive terms are attributed mostly to atmospheric path radiance, however the sensor calibration offset is commonly the larger term in the infrared wavelengths. The skylight term is only significant for shaded pixels (Kowalik et al., 1983), where ratio values are crude anyway due to relatively large signal quantization errors.

Several methods are available for estimating the combined atmospheric and sensor-offset correction term, including Dark-Pixel Subtraction (Crane, 1971), Radiance-to-Reflectance Conversion (Honey et al., 1974), regression-type techniques (Potter and Mendlowitz, 1975; Switzer et al., 1981; Crippen, 1986), and a statistical or

visual iterative-ratioing procedure (Crippen, 1986).

Examples of raw ratios (and ratios for which the additive terms have been incompletely removed) are plentiful in the literature. They are readily recognized by their significant retention (sometimes reversal) of topographic effects (such as in Figure 2c). The interpretability problems of raw ratios are easily explained by graphic and pictoral illustration.

Figure 3 is a simplified bi-spectral plot depicting two contrasting ground cover types in rugged terrain. Data adjustments for the additive-radiance terms have not been made, thus the plots do not project through (and intersect at) the graph origin. Ratio values increase circularly on the plot in the counter-clockwise direction. Although the lines represent spectrally homogeneous materials, ratio values along each of them are greater for pixels of low illumination intensity. Also, dark pixels of material U have the same ratio values as brighter pixels of material M, and are therefore indistinguishable. Note that this is not the case when the additive terms (45 and 6 for bands 1 and 4 respectively) are removed, as in Figure 1. Note also that these plots are derived from the image data of Figures 2a and 2b and explain the topographic reversal and spectral ambiguity of Figure 2c. M is freshly exposed bedrock and mining talus (high TM 1/4 ratio) and U is the adjacent undisturbed bedrock with coatings of desert varnish and other products of weathering (low TM 1/4 ratio).

Problems attributable to the use of raw ratios have been cited by Krohn et al. (1978) and Kowalik et al. (1983) in attempts to discriminate limonitic outcrops from non-limonitic outcrops. In other studies this cause of feature discrimination problems has gone unrecognized by the authors.

Some researchers have argued for the use of raw ratios because retained topographic information can provide benefits in visual image analyses. Most, however, have not acknowledged that these benefits <u>must</u> come at the expense of spectral distinctiveness. Others have argued that raw ratios have a greater signal-to-noise ratio than additive-term-adjusted ratios, however this is only true if one defines atmospheric path radiance to be "signal". For most applications it is more reasonably defined to be environmental noise. Clearly, removal of the additive terms is essential if ratio images are to coherently depict surface reflectance properties.

SIMPLE RATIOS: PROBLEMS

Ratios produced by simple division of additive-term-adjusted radiance data are closely related to the reflection properties of surface materials, and are herein termed "simple ratios". The formula for simple ratios is

$$V = ((R_n/R_d) * S) - C \qquad\qquad [4]$$

where V is the ratio value, Rn and Rd are the pixel values
of two bands, S is a scaling factor that spreads the results
over 256 quantization levels (with minor saturation), and C
is an offset term to position those levels at 0 to 255. For
general purpose use, problems remain for simple ratios.

One problem is that simple ratios show an awkward
complexity in that greater contrast is given to surface
materials that happen to have greater data values in the
band selected to be the numerator. This is because simple
ratios less than unity (numerator less than denominator)
have a range of zero to one (before scaling), while those
greater than unity have a range essentially from one to
infinity. A related problem is that while simple ratio
outputs greater than unity are distributed linearly (2/1,
3/1, 4/1, ...), their equivalent inverses less than unity
are distributed hyperbolically (1/2, 1/3, 1/4, ...).

Another problem with simple ratios is that their
distributions are commonly sharply peaked and the peak can
only be finely quantized if the output scale is increased to
the point where many non-peak values are saturated at the
range minimum and maximum.

As will be shown below, solutions to these problems are
not difficult.

BAND RATIOING FORMULAS

A wide variety of formulas can be used to produce ratio images that possess all of the previously stated desireable characteristics and avoid or reduce the problems of simple ratios. Several are presented below. In each it will be assumed that the data have been corrected for the additive terms. The following symbols are used:

V = Ratio value

R_n = Numerator-band pixel value

R_d = Denominator-band pixel value

P = Proportionality factor (constant)

S = Scaling factor (constant)

Except where noted, all of the ratio formulas discussed below are invertible and provide a generally advantageous quantization that relatively compresses (crudely quantizes) extreme values and expands (finely quantizes) values near unity. Since between-band differences in the overall magnitudes of band values have no environmental meaning (being a function of sensor gain settings), any ratio value can be reset to unity to take advantage of the fine quantization there. Commonly, the user will want to increase overall pixel discriminability (image entropy) by positioning the ratio modal value at unity (many ratio images are distinctly unimodal). $\overline{R_n}/\overline{R_d}$ (the ratio of the average additive-term-adjusted radiances) is a good first estimate of the ratio mean and mode values. Its inverse,

Rd/Rn, is therefore useful as a proportionality factor (P)
for centering the ratio values approximately at unity before
non-linear quantization. By selecting an appropriate
proportionality factor, any range of ratio values can be
emphasized in the quantization. Again, the proportionality
factor is just the fractional inverse of a selected value in
that range.

The scaling factor, S, is determined by making a low
guess, producing the ratio image and its histogram, and then
increasing S so that the output ratio image values fill the
quantization range with only minor saturation at the minimum
(0) and maximum (255) values. If saturation is (or projects
to be) one-sided and excessive then the midrange constant
representing unity (initially 128) can be adjusted to
reposition the ratio distribution. Only one trial iteration
should be necessary, and it may be economically beneficial
to grid-sample the image for producing the trial histogram.

Note that ratios having quantization stretches that are
symmetrical around unity, can be inverted:

$$\frac{\text{band 1}}{\text{band 2}} \quad \longleftrightarrow \quad \frac{\text{band 2}}{\text{band 1}}$$

by linear reversal:

        output value = range maximum - input value

when unity is assigned to the midrange value. Thus, such
ratios inverted on a display device (by reversing the lookup

table) are numerically equivalent to the same ratios
calculated with the numerator and denominator bands
interchanged.

All ratioing formulas presented will be evaluated and
compared by ratio modelling later in this paper and are only
briefly described here.

Log Ratios

Logarithmic ratios, first introduced by Goetz et al.
(1975), are derived by use of Equation 5.

$$V = 128 + \left(\left(\log \left(\frac{Rn}{Rd} * P \right) \right) * S \right) \quad [5]$$

Note that log ratios are inherently invertible around
unity because logs of ratios equal the negatives of the logs
of inverted ratios.

Arctangent Ratios

Arctangent ratios were first proposed by Wecksung and
Breedlove (1977) as another solution to the quantization
problems of simple ratios. Equation 6 is a practical
formula for their derivation.

$$V = 128 + \left(\left(atn \left(\frac{Rn}{Rd} * P \right) - 45 \right) * S \right) \quad [6]$$

The resultant values are linearly related to the

angular position of the pixels as plotted on a bi-spectral graph (after proportionality-factor adjustment).

Normalized-Difference Ratios

Normalized-difference ratios (Smedes et al., 1971) are very similar to log and arctangent ratios, yet are derived without the use of a complex function (Equation 7).

$$V = 128 + \left(\left(\frac{(Rn - Rd)}{(Rn + Rd)} \right) * S \right) \qquad [7]$$

In this case, however, to assure that the greatest stretch is applied to the selected ratio value, one band must be adjusted relative to the other before division occurs. This is achieved by substituting Rn*P for Rn in the formula.

The so-called Transformed Vegetation Index (TVI) is a modification of a normalized-difference ratio that uses a red band, Rn, and a near-infrared band, Rd (Rouse et al., 1974). The formula is:

$$V = \left(\left(SQR \left(\left(\frac{(Rn - Rd)}{(Rn + Rd)} \right) + K \right) \right) * S \right) + C \qquad [8]$$

where K=1 is used to avoid taking the square root of a negative number (constants less than 1 can be used but cannot be pre-determined to work for all pixel values). C is negative and is used to position the values between 0 and 255 after scaling. With K=1, the square-root function

asymmetrically compresses the normalized-difference values
toward unity, Rn/Rd=1 (Rn/Rd < 1 becoming larger and Rn/Rd >
1 becoming smaller). If K=/=1, however, then the
compression is toward some value other than unity (for the
commonly used K=0.5, compression is toward Rn/Rd = 3). No
logical quantization advantage is apparent for TVI and
several authors have reported no practical benefit over
simple ratios for measuring vegetation (e.g. Tucker, 1979;
Perry and Lautenschlager, 1984; Yool et al., 1986). Its
continued use therefore appears unjustified.

A square-root function can be forced to be symmetric,
to take advantage of it exponential quantization for general
purpose ratioing, by using a dichotomous formula (Equations
9a and 9b).

If Rn/Rd >= 1 then:

$$R = 128 + \left(\left(\left(SQR\left(\frac{Rn - Rd}{Rn + Rd} + 1\right)\right) - 1\right) * S\right) \quad [9a]$$

If Rn/Rd < 1 then:

$$R = 128 - \left(\left(\left(SQR\left(\frac{Rd - Rn}{Rd + Rn} + 1\right)\right) - 1\right) * S\right) \quad [9b]$$

It will be shown later by modelling, however, that the
square-root function has little effect on the quantization
of normalized-difference ratios.

Equilateral and Hyperbolic Ratios

When only the extreme values of a ratio image are of particular interest, it is beneficial to avoid compressed far-range quantizations. Equilateral ratios (Equations 10a and 10b) provide a quantization that is symmetrical around unity and is equal to the greater-than-unity quantization of simple ratios.

If $(Rn/Rd)*P >= 1$ then:

$$R = 128 + \left(\left(\left(\frac{Rn}{Rd} * P \right) - 1.00 \right) * S \right) \qquad [10a]$$

If $(Rn/Rd)*P < 1$ then:

$$V = 128 - \left(\left(\left(\frac{Rn}{Rd} * P \right) - 1.00 \right) * S \right) \qquad [10b]$$

By exchanging the if-then conditional statements, the quantization is still symmetrical but is equal to the relatively hyperbolic, <u>less</u>-than-unity quantization of simple ratios. In extreme contrast to equilateral ratios, this quantization is even more compressed in the far ranges than that of any other ratio discussed above. These quantization differences between equilateral and hyperbolic ratios are directly related to the great asymmetry of simple ratios. See the Modelling section below for histogram depictions of these quantization differences.

DEALING WITH VERY DARK PIXELS

A problem that can arise in ratios of any form is a light and dark, "salt and pepper" effect (figure 2d) in areas of very low terrain illumination (topographic shadows). This results where sensor noise, signal quantization noise, and ratio quantization noise combine to become large relative to the recorded signals. Thus, in areas of low illumination (areas of weak signal), ratio values can be so crude as to be virtually meaningless. Removal of these pixels from the ratioing process may be desirable. This can be done by assigning them a "neutral" value (perhaps 127 or 128), or a value chosen to be readily distinguishable in the output image (e.g., zero). The cutoff value for acceptible versus unacceptibly dark pixels is sensor, scene, and feature dependent, in addition to being a matter of judgement. In general, however, ratios using adjusted radiance values of less than 10 will produce results of questionable quality.

As a minimum requirement, pixels with band values of zero or less must be assigned some positive value before ratioing to avoid the mathematical impossibilities of division by zero or the taking of a log of a negative number. (Values of less than zero are possible after removal of the additive term when (1) a pixel is in deep shadow and the additive-term correction value for the scene is excessive for that specific pixel, or (2) when sensor noise results in a falsely low recorded radiance value.)

DISCUSSION

The need for using invertible-ratio formulas is
directly related to the magnitude of the range of ratio
values. When the range is small, then the differences
between the scale of simple ratios and the scale of their
inverses is small. Note, for example, that 1.1/1.0 and
1.0/1.1 are nearly equidistant from 1/1, but that 4/1 and
1/4 are not. Thus, invertible ratios are especially useful
for removing scale distortion when utilizing bands that are
poorly correlated. In general, these are also the bands
that provide the most information when ratioed.

In lithologic studies using Landsat MSS data, simple
ratios have generally been an adequate ratio form because
all bands are quite highly correlated. With Landsat TM data
and data from broad-spectrum airborne scanners this is
commonly not true, and the benefits of invertible ratios can
be substantial.

Non-linear ratio quantizations are generally beneficial
because many ratio-image histograms are strongly unimodal.
The non-linear quantizations described above compress the
distribution of ratio values in the extremes, and expand it
near unity. Thus, if the mode is positioned at unity, the
quantization is finer where most of the results lie, and the
average discriminability of each pixel from all other pixels
is increased (image entropy is increased). This is not
equivalent to distribution stretches applied after the

ratioing procedure, which can only reduce discriminability.
When the user is not concerned about overall
discriminability but instead wants to increase the spectral
discriminability of a selected range of ratio values, then
non-linear quantization is again beneficial when the
proportionality factor is properly chosen.

The differences between various types of ratios will
usually be more significant statistically than visually.
For statistical use (such as an input channel to a
classification algorithm), any increase in pixel
discriminability is potentially beneficial. For viewing,
however, pixel discriminability is subject to perceptual
constraints. The number of differing grey levels that can
be perceived in a scene is severely limited, varying
inversely with scene complexity from about 8 to more than
100. For a spatially complex ratio image, 256 grey levels
are an amount of detail far beyond what can be perceived,
therefore just how those levels are used is not very
important. The impact of quantization type on perception
should not be ignored, however, since the signal
(noise-free) components of ratio images generally do not
have high spatial complexity and, therefore the potential
exists to perceive several grey levels. Also, since users
commonly interactively stretch narrow ranges of ratio values
in displays to emphasize the detail of selected features, it
is beneficial to have superior quantization detail
available.

MODELLING OF THE RATIO OUTPUTS

In order to demonstrate the differences between the various types of ratios, a hypothetical ratio image was created with the ratio values distributed about a central value of 1/1 (unity). This modelling was necessary in order to demonstrate effects of the differing ratioing methods independent of any scene-specific features. The distribution of ratio values used for the modelling is arc-normal. That is, if the pixel values are plotted on a bi-spectral graph and their values are taken as the slopes (in degrees or radians) of the lines connecting them to the graph origin, then the histogram of those values has a gaussian distribution. Real ratio images commonly have distributions not unlike those created by this modelling. That is, when expressed in terms of simple ratios (instead of arcs), their distributions are unimodal and distinctly peaked.

The differences in the ratio outputs are easily seen in their histograms. All outputs were scaled to have 2% saturation equally split between the range ends.

In Figure 4a, simple ratio values are not distributed symmetrically and are sharply peaked below unity. The output image is therefore not invertible, and much information (in terms of the distinguishability of pixels) is lost.

Equilateral ratio values are symmetrical about the

central value, and the ratio image is therefore invertible. However, the histogram is sharply peaked such that several pixels are grouped and indistinguishable (image entropy is low). On the other hand, values significantly different from unity have a finely resolved quantization.

Hyperbolic ratio values are also symmetrical and invertible, but show a quantization stretch that is so severe at unity that it splits the arc-normal model histogram into a bi-modal distribution. Such a stretch could be advantageous for ratio images that are more peaked than the arc-normal model.

In Figure 4b, normalized-difference ratios, TVI ratios, and square-root normalized-difference ratios are shown to be similar, except for the non-invertibility of TVI. Note that the square-root function has only a minor effect on the peakedness of normalized-difference ratios.

In Figure 4c, arctangent, normalized-difference, and log ratios all display flattened distributions that are symmetrical (and therefore invertible) around their central value. These ratios are clearly superior to simple ratios in maximizing overall pixel discriminability, given the peakedness common in ratio image histograms.

CONCLUSIONS

Procedures are not complicated for maximizing the utility and overall information content of band-ratio

images. Simply put, radiance values need to be calibrated to ground reflectance, and then ratio values need to be optimally distributed within their quantization range to maximize pixel discriminability. For general purpose ratioing, invertible quantization is preferable so that the resultant image information is not dependent upon the assignment of the bands to the numerator or denominator.

This paper has reviewed the logical basis of band ratioing, discussed the desireable characteristics of ideal general-purpose ratio images, and described several potentially-useful ratioing formulas. Arctangent, normalized-difference, and log ratios are very similar, are easy to employ, and are generally preferred over simple ratios due to their invertibility and commonly-advantageous non-linear quantization of the resultant values.

ACKNOWLEDGEMENTS

NASA

REFERENCES

Crane, R.B., 1971, Preprocessing techniques to reduce atmospheric and sensor variability in multispectral scanner data: Proceedings of the Seventh International Symposium on Remote Sensing of Environment, p. 1345-1355.

Crippen, Robert E., 1986, The Regression Intersection Method of adjusting image data for band ratioing: ~~submitted to~~ International Journal of Remote Sensing, in press.

Crippen, Robert E., John E. Estes, and Earl J. Hajic, 1987, Band and ratio selection to maximize information content in color composite displays (in preparation).

Goetz, A.F.H., F.C. Billingsly, A.R. Gillespie, M.J. Abrams, R.L. Squires, E.M. Shoemaker, I. Lucchita, and D.P. Elston, 1975, Applications of ERTS images and image processing to regional geologic problems and geologic mapping in northern Arizona: Jet Propulsion Laboratory Technical Report 32-1597, 188p.

Honey, F.R., A. Prelat, and R.J.P. Lyon, 1974, STANSORT - Stanford Remote Sensing Laboratory pattern recognition and classification system: Proceedings of the Ninth International Symposium on Remote Sensing of Environment, p. 897-905.

Jackson, R.D., P.N. Slater, and P.J. Pinter, Jr., 1983, Discrimination of growth and water stress in wheat by various vegetation indices through clear and turbid atmospheres: Remote Sensing of Environment, v. 13, no. 3, p. 187-208.

Krohn, M.D., M.J. Abrams, and L.C. Rowan, 1978, Discrimination of Hydrothermally Altered Rocks Along the Battle Mountain- Eureka Nevada Mineral Belt Using Landsat Images: U.S. Geological Survey Open-File Report 78-585.

Knepper, Daniel H., Jr., and Gary L. Raines, 1985, Determining stretch parameters for lithologic discrimination on Landsat MSS band-ratio images: Photogrammetric Engineering and Remote Sensing, v. 51, no. 1, p. 63-70.

Kowalik, W.S., R.J.P. Lyon, and P. Switzer, 1983, The effects of additive radiance terms on ratios of Landsat data: Photogrammetric Engineering and Remote Sensing, v. 49, no. 5, p. 659-669.

Kreigler, F.J., W.A. Malila, R.F. Nalepka, and W. Richardson, 1969, Preprocessing transformations and their effects on multispectral recognition: Proceedings of the Sixth International Symposium on Remote Sensing of Environment, Environmental Research Institute of Michigan, p. 97-131.

Perry, Charles R., and Lyle F. Lautenschlager, 1984,
Functional equivalence of spectral vegetation indices:
Remote Sensing of Environment, v. 14, p. 169-182.

Maxwell, Eugene L., 1976, Multivariate systems analysis of
multispectral imagery: Photgrammetric Engineering and
Remote Sensing, v.42, no. 9, p. 1173-1186.

Podwysocki, M.H., D.B. Segal, and M.J. Abrams, 1983, Use of
multispectral scanner images for assessment of hydrothermal
alteration in the Marysvale, Utah, mining area: Economic
Geology, v. 78, no. 4, p. 675-687.

Potter, J.F., and M.A. Mendlowitz, 1975, On the
determination of haze levels from Landsat data: Proceedings
of the Tenth International Symposium on Remote Sensing of
Environment, p. 695-703.

Raines, Gary L., Terry W. Offield, and Elmer S. Santos,
1978, Remote-sensing and subsurface definition of facies and
structure related to uranium deposits, Powder River Basin,
Wyoming: Economic Geology, v. 73, no. 8, p. 1706-1727.

Rouse, J.W., Jr., R.H. Haas, J.A. Schell, and D.W. Deering,
1974, Monitoring vegetation systems in the Great Plains with
ERTS: Third Earth Resources Technology Satellite-1
Symposium, NASA SP-351, p. 309-317.

Rowan, L.C., P.H. Wetlaufer, A.F.H. Goetz, F.C. Billingsly, and J.H. Stewart, 1974, Discrimination of hydrothermally altered areas and of rock types using computer enhanced ERTS images, south central Nevada: U.S. Geological Survey Professional Paper 883, 35 p.

Rowan, Lawrence C., Alexander F.H. Goetz, and Roger P. Ashley, 1977, Discrimination of hydrothermally altered and unaltered rocks in visible and near infrared multispectral images: Geophysics, v. 42, no. 3, p. 522-535.

Rowan, L.C., and A.B. Kahle, 1982, Evaluation of 0.46- to 2.36-um multispectral scanner images of the East Tintic Mining District, Utah, for mapping hydrothermally altered rocks: Economic Geology, v. 77, no. 2, p. 441-452.

Switzer, Paul, William S. Kowalik, and Ronald J.P. Lyon, 1981, Estimation of atmospheric path-radiance by the Covariance Matrix Method: Photogrammetric Engineering and Remote Sensing, v. 47, no. 10, p. 1469-1476.

Tucker, Compton J., 1979, Red and photographic infrared linear combinations for monitoring vegetation: Remote Sensing of Environment, v. 8, no. 2, p. 127-150.

Vincent, Robert K., 1972, An ERTS Multispectral Scanner experiment for mapping iron compounds: Proceedings of the Eighth International Symposium on Remote Sensing of Environment, p. 1239-1247.

Wecksung, G.W., and J.R. Breedlove Jr., 1977, Some techniques for digital processing, display and interpretation of ratio images in multispectral remote sensing: Society of Photo-Optical Instrumentation Engineers, v. 119, Applications of Digital Image Processing, p. 47-54.

Yool, Stephen R., Jeffrey L. Star, John E. Estes, Daniel B. Botkin, David W. Eckhardt, and Frank W. Davis, 1986, Performance analysis of image processing algorithms for classification of natural vegetation in the mountains of Southern California: International Journal of Remote Sensing, v. 7, p. 683-702.
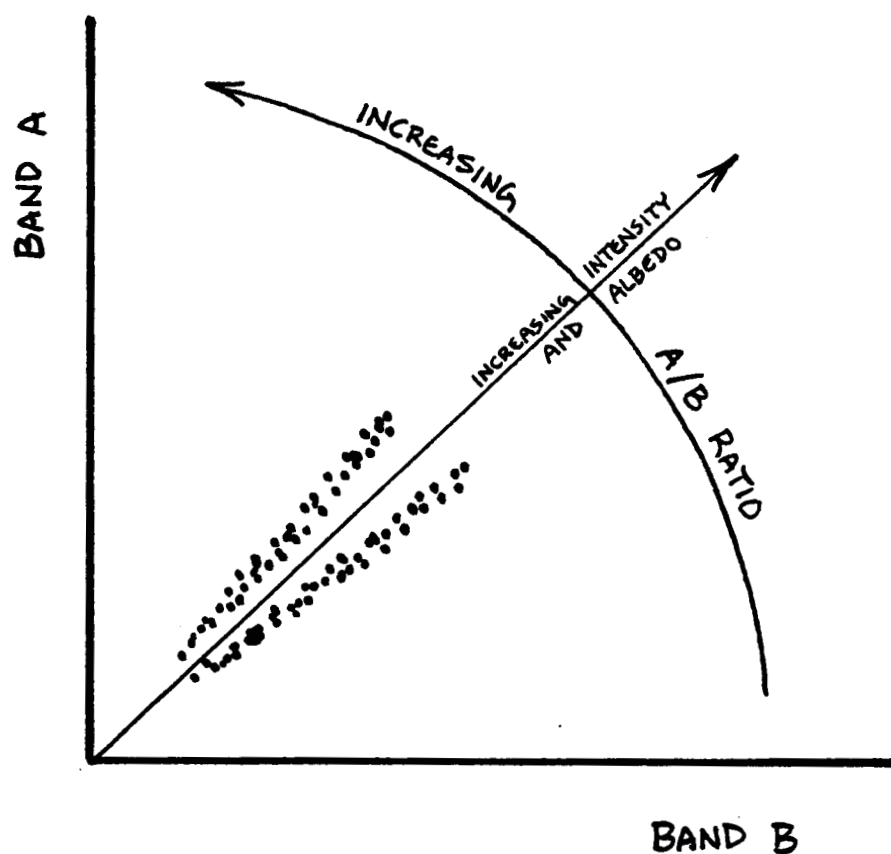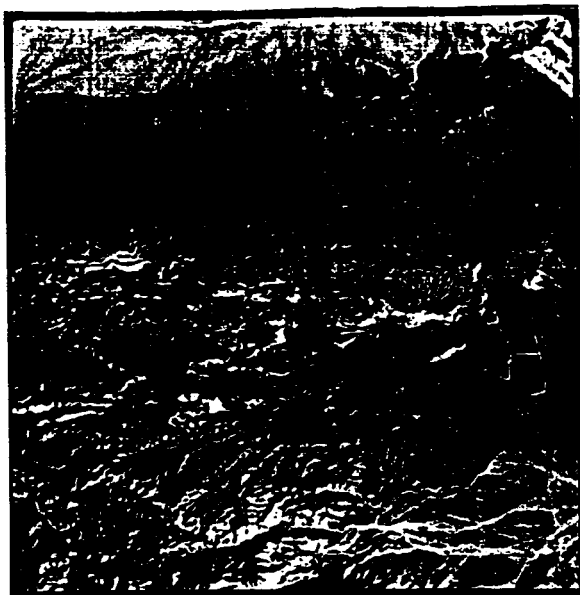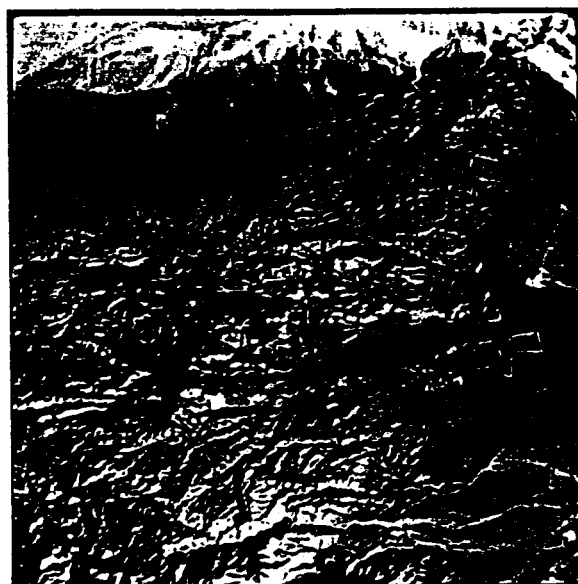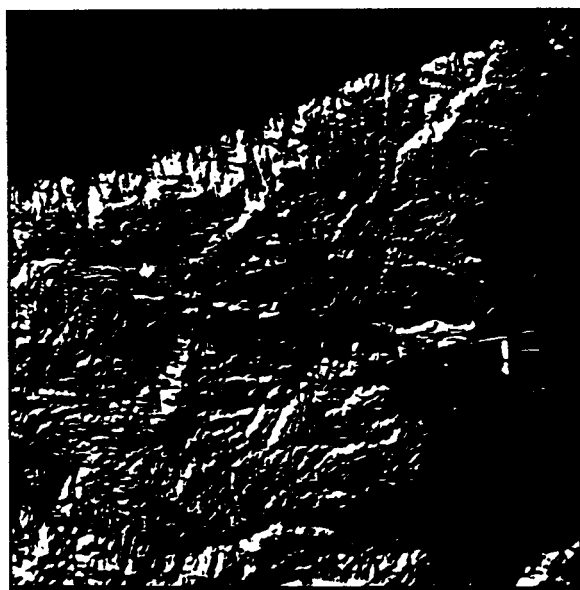
Figure 1. Bi-spectral plot of two surface materials of differing
reflectance properties in rugged terrain. Pixel values are
calibrated so that zero in each band represents zero ground
radiance. The materials are readily distinguished by their
ratio values despite having pixel values that greatly overlap
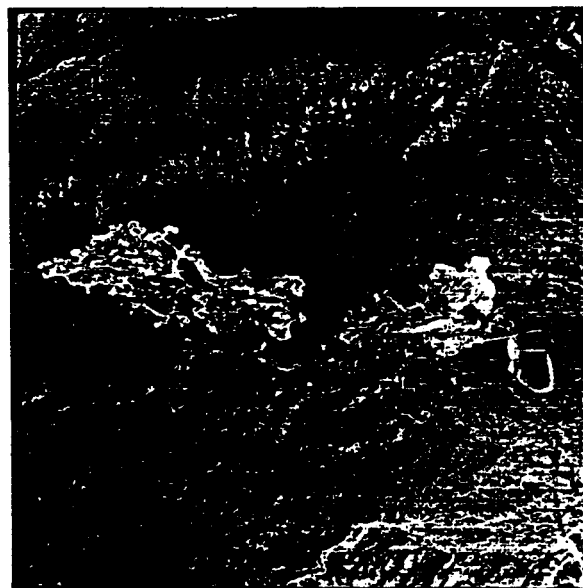in the individual bands.

Figure 2. Landsat TM imagery of the Eagle Mountain Iron
Mine and vicinity, Southern California, 12 December 1982.
(a) Band 1. (b) Band 4. (c) Raw ratio 1/4 showing
reversal of topographic expression. (d) Log ratio 1/4 with
additive-radiance terms removed. Freshly exposed bedrock
and talus are clearly depicted as bright areas in d, but
remain confused with adjacent areas in c due to the
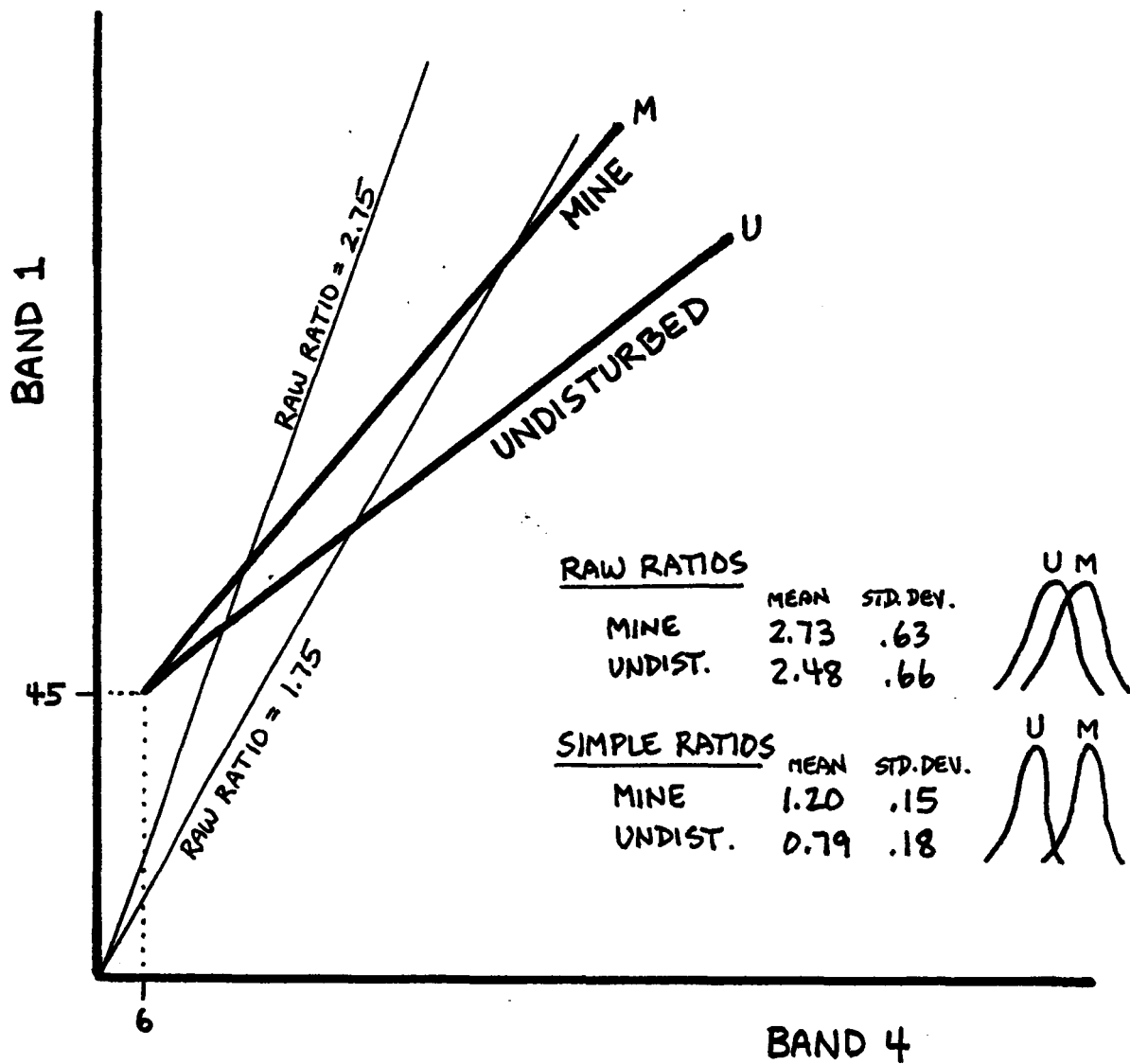preservation of topographic effects.

Figure 3. Simplified bi-spectral plot of image data from Figure 2.
Raw ratios do not differentiate the surface materials since
their values are equal for brightly illuminated mine pixels
and darker pixels of undisturbed ground. Also, dark pixels
of both surface materials tend to have greater raw ratio values
than brightly illuminated pixels, resulting in the reversal
of the topographic effect as seen in Figure 2c. Removal of
the additive terms (45 and 6) results in distinct ratio values
for the two spectrally distinct materials (as in Figures 1 and 2d).
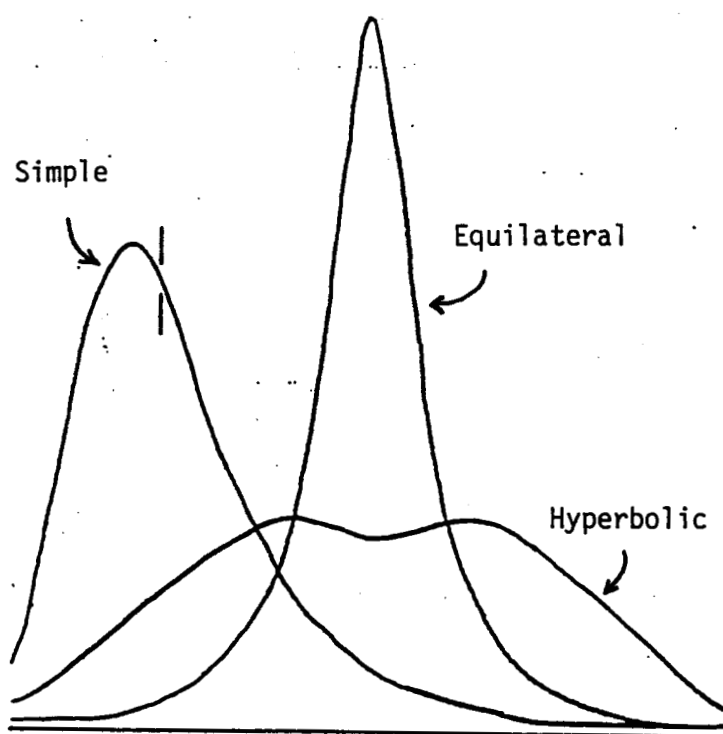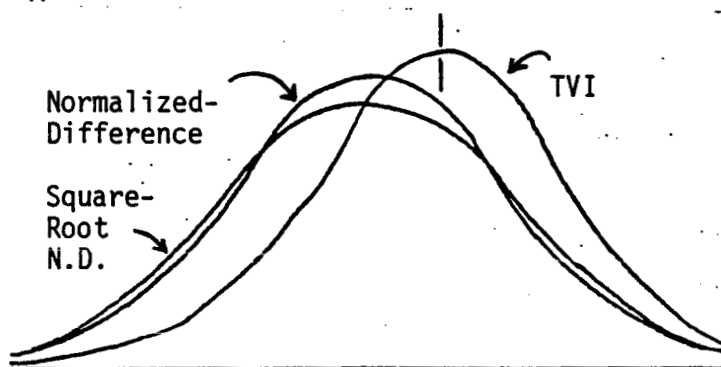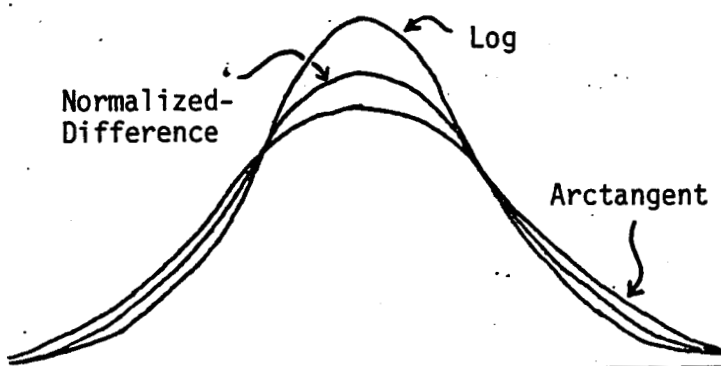
Figure 4. Histograms of arc-normal modelled
ratio image. Unity is at the mid-range value for
all symmetric (invertible) ratios and is indicated
by a vertical line for simple ratios and TVI.

Glossary of Terms and Abbreviations
Commonly Used in Remote Sensing Research and Applications

Jeffrey L. Star, John E. Estes, Kristi Lombard

Glossary of Terms and Abbreviations
Commonly Used in Remote Sensing Research and Applications

Jeffrey L. Star, John E. Estes, Kristi Lombard

The jargon and technology of remote sensing research and applications move rapidly. One of the problems we have discovered in recent workshop meetings is a difference in the way people are using a variety of terms. The following list of terms and abbreviations is culled from a variety of sources, principally in non-referreed sources.

In the next year, we propose to continue circulate this list among a number of our NASA-sponsored colleagues. The intent of this activity is to develop a concensus definition of the terms, to further communication within the profession.

AAG - Association of American Geographers.

Accuracy - A measure of freedom from error.

Active System - In remote sensing, a system that is the source of
the electromagnetic radiation reflected or scattered by the
object being sensed. For example, radar is an active microwave
system (compare with "passive system").

AGIS - Automated Geographic Information System.

AgRISTARS - Agriculture and Resources Inventory Surveys Through
Aerospace Remote Sensing.

AGU - American Geophysical Union.

AIAA - American Institute of Aeronautics and Astronautics.

AID - Agency for International Development (U.S. federal agency).

Algorithm - A procedure for performing a specific action; in
software this is a set of instructions to the computer.

Alphanumeric - Consisting of both letters and numbers.

Alphanumeric and graphics terminals - Usually a black and white
TV screen with attached keyboard, or a keyboard terminal with a
graphics printer, for interacting with a computer.

Analog - the representation of numerical quantities by means of
physical variables (voltage, current, etc.) as opposed to
"digital".

Analog Computer - A calculating device that operates on numbers
represented as measurable quantities (e.g., voltages) rather than
the numbers themselves (compare with "digital computer").

Analyzed Data Records (ADRs) - Those records which the
investigator designates as the best to display the scientific
results of an experiment and provide the physical quantities by
applying calibration curves or algorithms to the corrected
observed quantities of the Reduced Data Records. The data may be
time averaged and may incorporate model-dependent assumptions to
obtain the physical quantities. Charts, graphs, table,
correlation coefficients, model parameters, photographs, and
plots are possible forms of these records.

Ancillary Data - Additional, supplemental data.

Angular Field of View - See Field of View (FOV).
Antenna - The device that radiates electromagnetic radiation from
a transmitter or receives radiation from other sources or
antennas.

GLOSSARY

Array Processor - A specialized computation unit used in connection with a regular computer CPU, which performs operations an arrays (matrices) of data. Used when data arrays must be manipulated at high speed.

ASCS - Agricultural Stabilization and Conservation Service (USDA)

ASP - American Society of Photogrammetry.

B/W - Black and White

Band - A range of wavelengths of electromagnetic radiation specified to produce a single response on a sensing device; also known as a channel.

Beam Width - A measure of the concentration of power by a directional antenna. It is the angle subtended at antenna by arbitrary (normally one-half the maximum) power-level points across the axis of the beam.

BIL - Band Interleaved by Line

Bit - In digital computing, a binary digit which may either be one or zero.

BPI - Bits Per Inch

BPS - Bits Per Second

BSQ - Band Sequential

Byte - A collection of 8 bits so that one byte can represent any number between 0 and 255.

Calibration - the act of comparing certain measurements in an instrument with a standard.

Cartography - The science of representing graphically the known physical features of the Earth's surface, whether natural or artificial.

Catalog - A collection of detailed information about whole data sets.

Cathode Ray Tube (CRT) - An electron tube whose face is covered by a phosphor that emits light when energized by its electron beam; for example, a computer display screen or a television.

CBD - Central Business District

CCRS - Canada Centre for Remote Sensing

Cell Based Data Structure (CBDS) - A logical organization of spatial data in which geographic space is modeled as a surface

composed of cells.

Census Tract - Small areas averaging 4000 in population and varying in land area. They represent neighborhoods having similar socioeconomic characteristics and are defined by local committees in cooperation with the Bureau of the Census.

Central Processing Unit (CPU) - That part of the computer hardware which performs the actual calculation and manipulation of data.

Channel - In remote sensing, a narrow spectral interval; also referred to as a band.

CIR - Color Infrared Film

Co-Investigator (Co-I) - An associate of the Principal Investigator (PI) who is assigned a supporting role in the investigation. In addition, some data rights may be assigned to the Co-I by the P-I.

Compatible - Executable on more than one computer.

Computer Compatible Tapes (CCTs) - Tapes containing data in machine-readable (digital) format. Four Landsat CCTs are required to cover one Landsat scene.

Contour Map - A topographic map that portrays relief by the use of contour lines of equal elevation.

Contrast - The amount of difference between the intensity of light areas and dark areas of an image. The highest contrast ratio is between black and white.

Contrast Enhancement - A mathematical procedure usually performed by a computer to artificially increase the contrast in an image, making certain features stand out sharply.

COSMIC - Computer Software Management and Information Center

CPU - Central Processing Unit

CRT - Cathode Ray Tube

CT - Census Tract

CTS - Communications Technology Satellite (Canadian)

CZCS - Coastal Zone Color Scanner

D/A - Digital to Analog

Data Set - A collection of similarly formatted records having like information (from one or multiple sensors), identical

processing, and, typically, common temporal/spatial/spectral resolution.

Data Set Granule - A segment of a data set, such as a single image, one or several records in some fixed time increment, or files on each of several tapes.

Data System - A package combining data directory, catalog, and/or inventory functions, as well as data access, sometimes manipulation, and display capabilities.

DEM - Digital Elevation Model

Digital Computer - A calculating device that operates on the principle of counting rather than measuring quantities that represent numbers (compare with "analog computer").

Digital Format - Data storage in machine-readable form.

Digitizer - A device that converts analog information into machine-readable form.

Digitizing - The process whereby an analog value such as position on a map is converted into digital coordinates or values.

DIME - Dual Independent Map Encoding

Directory - A collection of high level information about whole data sets (names, locations, sources of further information, other items TBD).

Discipline Issues - Those SD issues which are associated with spatial data characteristics (see Section 2.1).

Disks - Large capacity, direct access storage devices; data is stored as magnetic areas on the face of one or several platters. These platters rotate rapidly and each bit of data is available to the reading/writing "head" once per revolution, providing high-speed data access. Also spelled discs.

DLG - Digital Line Graph

DMA - Defense Mapping Agency (U.S. Department of Defense)

DN - Digital Number

DSN - Deep Space Network

EDC - EROS Data Center (USGS, Sioux, Falls, SD)

EDP - Electronic Data Processing

ELAS - Earth Resources Laboratory Application Software

GLOSSARY

Electromagnetic Radiation - Energy propagated through space or a material in the form of waves of interacting electric and magnetic fields; for example, radio waves and light.

Electromagnetic Spectrum - An ordered array of known electromagnetic radiations including ultraviolet radiation, visible radiation, infrared radiation, and microwaves (see Figure 2-1).

EMR - Electromagnetic Radiation

Entity Based Data Structure (EBDS) - A logical organization of spatial data in which geographic space is modeled as a surface composed of spatial entities described by locational identifiers related to a coordinate origin.

EOF - End of File

EOT - End of Tape

EPA - Environmental Protection Agency

ERIM - Environmental Research Institute of Michigan

ERL - Earth Resources Laboratory (NASA, NSTL Station, MI)

EROS - Earth Resources Observation Systems (USGS, Sioux Falls, SD)

ERAS - Earth Resources Technology Satellite (now called Landsat)

ESA - European Space Agency

Experiment - A term used interchangeably with investigation (the latter is preferred). Activity or effort aimed at the generation of data obtained by measurement of space phenomena or the use of space to observe earth phenomena and the resulting analysis of such data.

Experiment Data Records (EDRs) - Those records provided to the Principal Investigator, Team Leader, Guest Investigator, Co-Investigator, or team member containing all the data from the mission required to carry out the investigation specified in the contract or launch agreement. These records may include orbital position, spacecraft attitude, instrument attitude, commands, housekeeping data, ground time, spacecraft time, data from other investigations and other information as agreed upon. The exact form of these records and manner in which these data are provided may vary depending upon the policies, procedures, and capabilities of the project, the payload or mission control centers, the data acquisition network, and any support processing facilities. These records shall be specified in the Project Data Management Plan.

GLOSSARY

Facility-Class Payload Mission - A mission in which the payload is an instrument or set of instruments which serve as a facility for a large group of Guest Investigators who may be selected at different times throughout the life of the mission to participate. This type of mission may not have Principal Investigators or Team Leaders and all the data collected from such a mission is generally maintained by the project for use by Guest Investigators. Availability of data for the scientific community at large shall be specified at the Project Data Management Plan.

False Color Composite - An image formed by assigning colors to two or more black and white images of the same scene and combining them in a computer or on film to form a single image. The resulting image pinpoints differences and similarities between the original images.

Field of View (FOV) - The solid angle through which an instrument is sensitive to radiation.

FL - Focal Length

Flux - The rate of flow of some physical quantity, usually energy.

Footprint - The area on the surface being investigated by a remote sensing device; it is approximately given by the product of the beam width (in radians) times the altitude of the remote sensing platform.

FORTRAN - A commonly used, high level algebraic language for coding computer algorithms (derived from FORmual TRANslation).

FOV - Field of View

FSK - Frequency Shift Key

Gamma - Unit of magnetic field strength equal to 0.00001 gauss (Earth's surface magnetic field strength is about 0.5 gauss or 50,000 gammas).

GBIS - Geo-Based Information System

Geostationary Satellite - An equatorial satellite circling the Earth at such a high altitude (about 36,000 km) that its period of revolution about the Earth matches the Earth's rotational period; the satellite appears fixed in the sky to an observer on the surface, and the surface of the Earth appears fixed to the satellite. Also called geosynchronous.

GIS - Geographic Information System

GMT - Greenwich Mean Time

GOES - Geostationary Operational Environmental Satellite

GRD - Ground Resolve Distance

Ground Truth - Information obtained on site; frequently used to assist and corroborate interpretation of remotely sensed data.

Guest Investigator (GI) - Investigator selected to conduct observations and obtain data within the capability of a NASA mission, which are additional to the mission's primary objectives, or for a facility-class payload mission.

H/W - Hardware

Hardware - The physical components of a computer.

HDDT - High Density Digital Tape

HDT - High Density Tape

HDTR - High Density Tape Recorder

Heat flux - The rate of flow of hear energy through an object.

High Level Programming Language - A programming language (such as FORTRAN) which is close to everyday, written language; a single line of code written in a high level language can represent dozens of machine instructions.

HOM - Hotline Oblique Mercator

Human Engineering - The design and implementation of systems taking human psychological concerns into consideration.

I/O - Input/Output

IBIS - Image Based Information System

IFOV - Instantaneous Field of View

IG - Initial Gap

Image - The representation of a scene as recorded by a remote sensing system (includes both machine-readable storage and photo products).

Image Classification - Analysis of digital image values, including spatial, temporal, and spectral band relationships, to obtain categories of pixels or specific features.

Image Display Terminal - Displays color and/or black and white images on a television screen. There may be an attached function keyboard to control manipulation of the image. Otherwise manipulation is performed via the computer to which the terminal is connected

GLOSSARY

In-code Documentation - The level at which comments have been added to the program's source code. A well-documented program is easier to modify or transport.

Information - Data that have been processed for a particular use.

Input - Information or data transferred or to be transferred from an external storage medium into the internal storage of the computer.

Input/Output Handler - The software which controls access from the CPU to peripheral devices, including storage units and display terminals. This software may be written for use by a single program, separate package of subroutines which are used by several programs, or a separate package supplied with the computer.

Interactive - A method of operation that allows instantaneous communication between man and machine.

Interoperability - Electronic interconnection of NCE modules, with appropriate supporting software, allowing sequential and concurrent exercise of system functionality in a quasi-homogeneous and convenient way, and allowing automatic migration of information among modules.

Interval - 1) The time between two events; 2) as a measurement scale, describes ranks and classes of data where the difference between the ranks is defined and constant, but may be arbitrary (e.g., Celsius and Fahrenheit temperature scales).

Inventory - A collection of information about the granules of a data set. (Note that inventories giving names, locations, and independent variable ranges of data set granules may need to be distinguished from those inventories characterizing data set granules by their contained data - e.g., % cloud coverage.)

Investigation - Activity or effort aimed at the generation of data obtained by measurement of space phenomena or the use of space to observe earth phenomena and the resulting analysis of such data.

Investigator - A participant in an investigation. This term may refer to a Principal Investigator, Co-Investigator, Team Leader, team member, Guest Investigator, or any other member of an investigation group.

IR - Infrared

IRG - Inter-record Gap

JPL - Jet Propulsion Lab (NASA, Pasadena, CA)

GLOSSARY

JSC - Johnson Space Center (NASA, Houston, TX)

KBPS - Kilobits Per Second

L/C - Land Cover

L/S - Landsat

L/U - Land Use

LACIE - Large Area Crop Inventory Experiment

Land Use - Human-Imposed functions of a land area.

Landcover - The vegetation and artificial constructions covering
a land surface.

LANDSAT - Land Satellite (Formally Called ERTS)

Landsat - A series of Earth-observing satellites that serve as
platforms for the RBV and MSS Instruments and will serve as a
platform for the Thematic Mapper.

Language Percentages - The fraction of a program or package which
is written in each of several languages.

Level of Man-Machine Interface - The existence or extent of
dialogue between the computer program and the user. The highest
level of interface consists of question/answer or menu-driven
programs; the lowest level of interface includes batch-submittal
of jobs via cards or card images.

LFC - Large Format Camera

LIDAR - Light Detection and Ranging

LL - Line by Line

LLA - Adjusted Line Length

Locational Identifier - A means of assigning content data to
spatial entities. Includes external descriptive identifiers,
such as county name, identifiers which assign content data to
location within an array of discrete units (e.g., pixels), and
coordinate identifiers related to an origin.

Low Level Programming Language - A programming language very
close to the actual computer instruction set; commonly called an
assembly language (compare with "High Level Programming
Language").

LP - Line Printer

LPM - Load Point Marker

Machine Language - The computer instruction set (usually represented just by numbers; if characters and numbers are used, this becomes assembly language).

Magnetometer - A device for measuring Magnetic fields.

MAGSAT - Magnetic Field Satellite (U.S.)

Manual Classification - The identification of features on aerial or satellite photographs by tone, color, texture, pattern, shape, and size.

Map - Usually a two dimensional representation of all or part of the Earth's surface, showing selected natural or man-made features or data, preferably constructed on a definite projection with a specified scale; includes digital maps, and other special maps.

Map Projection - See Projection.

MAPSAT - Mapping Satellite (U.S.)

Maximum Likelihood Rule - A statistical decision criterion to assist in the classification of overlapping signatures; overlapping units are assigned to the class of highest probability.

MCR - Monitor Console Routine

Measurement Scale - A system for quantifying observations according to predetermined rules, which define four successfully greater levels of data precision (nominal, ordinal, interval, and ratio); see Section 2.2.7.

Medium - Physical device on which data are stored, such as a photographic product, map, table, card, or magnetic tape.

Memory - An organization of data storage units in the computer; also the amount of main core storage required for a computer task, provided by direct access via the central processing unit.

Mission - One or more flights within an approved payload project.

Mission Scientist - A scientist from a NASA field center assigned to a Spacelab mission, the Mission Scientist has similar functions as the Project Scientist with the exception of direct responsibility for the development of any experiments.

MLA - Multi-Linear Array

MPP - Massively Parallel Processor - GSPC

MSS - Multispectral Scanner

MTU - Magnetic Tape Unit

Multispectral (line) Scanner (MSS) - A remote sensing device capable of recording data in the visible and infrared portions of the spectrum.

Nadir - The point on the ground vertically beneath the center of the remote sensing system; also called the sub-satellite point.

NASA - National Aeronautics and Space Administration

NASA Catalog Environment (NCE) - The aggregation of the data information modules (directories, catalogs, inventories) of NSSDC, project- and discipline-specific systems, and of individual scientists; the interfaces to, and links among, these modules; the use and management of these modules; etc. (Note that NCE is viewed more as a coordinated grouping of systems rather than as a single supersystem.

NASCOM - NASA Communications Network

National Space Science Data Center (NSSDC) - The main central repository for selected data and documentation from space science flight missions that serves as a disseminator of this archived data and supporting information to users throughout the international scientific community. The NSSDC, located at Goddard Space Flight Center, serves as a switching center for requesters who desire data still held individually by Principal Investigators (PIs) or Team Leaders (TLs) by providing a description of the spacecraft and experiment and the name, address and telephone number of the PI or TL. For missions involving a Guest Investigator program in association with a PI or TL experiment or involving a facility-class payload the role of the NSSDC shall be specified in the Project Data Management Plan.

NCIC - National Cartographic Information Center (USGS)

Nimbus - A series of Earth-observing experimental weather satellites carrying a variety of sensors; the last of the series (Nimbus-7) was launched late in 1978.

NMD - National Mapping Division (USGS)

NOAA - National Oceanic and Atmospheric Administration

Nominal - As a measurement scale, distinguishes things only on the basis of their intrinsic character (difference in kind, e.g., apples, oranges).

NOS - National Ocean Survey

NRZ - Non-Return to Zero

NSF - National Science Foundation

GLOSSARY

NSTL - National Space Technology Laboratory (Bay St. Louis, MI)

NTIS - National Technical Information Service

OCS - Ocean Color Scanner

Off-line - Peripheral equipment that is not under the control of the CPU, as contrasted to on-line.

Off-nadir Viewing - Intercepting electromagnetic radiation along a direction other than the vertical (nadir) between the sensor and the object being sensed (normally the Earth).

OMB - Office of Management and Budget

On-line - Under direct control of the Central Processing Unit.

Operating System - The high-level administrative program running in a computer at all times which controls the overall operation of the computer and its tasks.

Ordinal - As a measurement scale, distinguishes things on the basis of rank by some quantitative measure (e.g., small, medium, large).

OSTA - Office of Space and Terrestrial Applications (NASA)

Output - Information, data, or other results of a computer operation which are recorded on some external storage device (tape, disk, printer, etc.).

Passive System - In remote sensing, a system that responds to electromagnetic radiation which is either emitted by the object or coming from a natural source, such as the sun, and is reflected or scattered by the object (compare with Active System).

Peripheral Device - A device connected to a computer to provide communication (e.g., alphanumeric terminal, printer, plotter) or auxiliary functions.

PI - Photo Interpretation

PI - Principal Investigator

Pixel - "Picture element": The smallest unit of surface reflectance measured by a sensor system.

Platform - In remote sensing, the physical object (e.g., balloon, rocket, or satellite) that carries the instrument (sensor) that makes the remote measurement.

Plotter - A printer used to record non-alphanumeric information, such as graphs or images, on paper.

GLOSSARY

PMT - Photomultiplier Tube

Precision - The degrees of exactness with which a quantity is stated; contrast with "accuracy," which refers to the absence of error, regardless of precision.

Preprocessing - Manipulation of raw data for standardization prior to further analysis.

Principal Investigator (PI) - A person who conceives an investigation and is responsible for carrying it out, reporting its results, and providing appropriately selected data and supporting documentation to the scientific community in accordance with the Project Data Management Plan. The PI chooses his Co-Investigators and assignes them roles and privileges. The PI is the primary point of contact with the project office regarding the investigation.

Printer - An output device used to record the output (usually alphanumeric) on paper.

Program Scientist - A NASA Headquarters official assigned to each mission who has a number of roles and responsibilities defined in NASA Management Instruction 7100.11, Attachment D. The most relevant one for this subpart is the responsibility to establish the data analysis, data dissemination, and data archiving policies for the mission, which will be documented in the Project Data Management Plan.

Project Data Management Plan (PDMP) - A plan that addresses the total activity associated with the data acquired by a mission from the delivery of the Experiment Data Records to the Investigators to the delivery of selected reduced and analyzed records along with supporting documentation to a specified repository. The plan should provide the milestones in the data reduction, data interpretation, and resource requirements for these phases. Any planned data interpretation meetings, workshops, or other activities should be identified. The type of data records, data products and compilations that have been selected in concert by the investigators, the Project Scientist, the NSSDC acquisition manager, and any appropriate scientific advisory personnel for general availability to the international scientific community and for delivery to a disseminating repository, such as the NSSDC, shall be specified. For missions where the data will be maintained for many years by the project, the Principal Investigator handling a Guest Investigator program, or by an Institute established by the mission, the eventual transfer of appropriate data to a more permanent archive, such as NSSDC or other repository, shall be specified. Conditions for discarding or destroying the Experiment Data Records shall be specified.

Project Scientist - A scientist from a NASA field center assigned to a project to manage the scientific aspects. The roles and responsibilities of this function are given in NASA Management

GLOSSARY

Instruction 7100.11, Attachment E.

Projection - A systematic construction of features on a plane surface to represent corresponding features on a spherical surface. These features include observable phenomena (e.g., physical and cultural features such as coastlines and highways) as well as constructs (e.g., lines to represent parallels and meridians, political boundaries, or statistical units).

Proximal Sensing - Obtaining information by making direct, physical contact with it (compare to "remote sensing"); also making direct or in-situ measurements.

PS - Polar Stereographic (Map Projection)

PSK - Phase Shift Keying

QA - Quality Assurance

R/S - Remote Sensing

R&D - Research and Development

Radiance - Flux of radiant (electromagnetic) energy measured in power units (e.g., watts); frequently confused with radiancy (flux density per solid angle, commonly measured in units of watts cm-2 sterad-1).

Radiometer - A passive device for intercepting and quantitatively measuring electromagnetic radiation in some bank of wavelength.

Radiometric Characteristics - Characteristics of a radiometer that help to define its resolution.

Ratio - A measurement scale; distinguishes things on the basis of magni-image line by line.

Ratio - As a measurement scale, distinguishes things on the basis of magnitudes that are intrinsically meaningful by use of a non-arbitrary zero point (e.g., age or Kelvin temperature scale).

Rawinsonde - An instrument balloon carrying temperature and humidity probes in addition to a radio which transmits the information to an Earth station.

RBV - Return Beam Vidicon

Rectification - The process of projecting an oblique or tilted image onto a horizontal plane to produce the equivalent of an untilted image.

Reduced Data Records (RDRs) - Those records prepared from the Experiment Data Records by applying corrections, where applicable, for temperature, voltage, gain change, offsets, dead time, drift and other known instrument changes, as well as

eliminating unusable noisy periods and periods of questionable instrument performance. The Reduced Data Record should contain all the basic and supporting measurements obtained from the experiment, such as time, position, attitude, settings of instrument by command, housekeeping data and other information needed to analyze the data in an independent fashion. Visual data, such as photographs derived from imaging processing techniques, may also be considered as RDRs.

Registration - Superposition of points on one image with corresponding points on a second image or map of the same scene.

Remote Sensing - Obtaining information about an object or phenomenon without direct contact. (Compare with Proximal Sensing.)

Resampling - The use of mathematical or geometrical methods to estimate values on one cell based structure from values originally given on another structure; may include interpolation and extrapolation.

Resolution - The minimum area of surface that can be imaged by a remote sensor system.

Return Beam Vidicon (RBV) - A modified vidicon television camera tube that produces high resolution images. On Landsat 1 and 2 there were three cameras operating in the green, red, and IR spectral regions to form the RBV system; Landsat 3 has two black and white cameras each covering half the swath width. These images are not used as frequently as the MSS images.

RF - Radio Frequency

RJE - Remote Job Entry

RT - Real Time

S/C - Spacecraft

S/W - Software

SAB - Space Applications Board (National Research Council)

SAR - Synthetic Aperture Radar

Scalar - A quantity with a numeric value.

Scale - The ratio of distance on a map, chart, or photographic image to the equivalent distance on the Earth's surface.

Scaling - Transforming a display by multiplying all dimensions by a constant value, magnifying or reducing features.

Scientific and Technical Information Facility (STIF) - NASA's document and report acquisition and abstracting facility that

produces a biweekly abstract journal, STAR, covering the
aerospace report literature and a biweekly abstract journal,
IAA, covering the published literature in these fields. The
facility also produces microfiche copies of the report
literature for primary distribution.

SEASAT - Sea Satellite (U.S.)

Seasat - An Earth-observing satellite designed to gather
information about the oceans; launched in 1978, died abruptly in
1978.

Sensor - A device that gathers electromagnetic radiation and
presents it in a form suitable for obtaining information about
the environment.

Signature - A set of spectral, tonal, or spatial characteristics
which together serve to identify a class or feature by remote
sensing.

SLAR - Side Looking Airborne Radar

SMS - Synchronous Meteorological Satellite

SMSA - Standard Metropolitan Statistical Area

SNR - Signal-to-Noise Ratio (also S/N)

Software - A computer program as written in a high or low level
language; may include the documentation explaining the program.

Software Organization - The manner in which the computer program
is arranged in memory. The simplest organization is one task or
several tasks which execute serially. The more sophisticated
arrangements include tasks which are linked through the operating
system or through a unique and self-contained linking program,
and a single task decomposed to segments, which communicate with
each other but execute independently one at a time (overlays).

Software Transportability - A rating given to the anticipated
ease with which a program could be moved to another computer.
This rating depends on the software's modularity, its man-machine
interfaces, its size and organization, and language in which it
is written, the use of I/O handlers, and the documentation
available.

Space Science Flight Investigations - Investigations of natural
phenomena of the earth and its environment, the moon, other
planets, the sun, interplanetary space, and other celestial
object and regions made from aircraft, balloons, sounding
rockets, satellites, probes, and manned spacecraft for the
purpose of increasing basic knowledge of these natural phenomena.
Biological investigations involving the search for
extraterrestrial life are included.

GLOSSARY

Spatial Data - Phenomena with implicit or explicit locational identifiers; includes spatial data content and spatial data entities.

Spatial Data Content - Direct or indirect measurements that are spatially distributed, which means that their values vary from one location to another, and data which relate to a single location, where variation in value occurs through time.

Spatial Data Entities - Points, lines, areas, or surfaces with which spatial data content are associated (e.g., weather stations, rivers, watersheds, and elevation models).

Spatial Data Integration (SDI) - The process of combining multiple spatial data sets and providing for their storage, retrieval, analysis, and display.

Spatial Data Structures - The logical organization of spatial data entities and spatial data content.

Spectrometer - A radiometer with a dispersive element (prism, grating, or circular interference filter) so that the incident radiation is measured as a function of wavelength.

SPOT - Satellite Probatoire d'Observation de la Terre (France)

Standard Metropolitan Statistical Area (SMSA) - A county or group of contiguous counties that contain at least one city of 50,000 innabitants or more, or "twin cities" with a combined population of at least 50,000. In addition to the county or counties containing such a city or cities, contiguous counties are included is an SMSA if, according to certain criteria, they are socially or economically integrated with the central city.

Supervised Classification - Classification of images using statistics developed from training sites.

Swath Width - The area on either side of a platform which is surveyed by a remote sensing instrument; composed of many (overlapping) footprints.

Tape - A long, thin, flexible medium for storing data; usually refers to digital magnetic tape on which data are stored as a string of ones and zeros by altering the magnetic domains of an iron oxide coating on the tape; less frequently refers to analog magnetic tape or occasionally to paper tape on which data are stored by punching holes through the paper.

Tape Drive - A device used for storing data on magnetic tape. Tape drives may record up to none tracks of data along the width of a standard computer tape.

TBD - To Be Determined

TDRSS - Tracking and Data Relay Satellite System

Thematic Mapper - A high speed, high resolution multispectral scanner for monitoring Earth resources; planned for Landsat-D.

TIRS - Thermal Infrared Scanner

TLM - Telemetry

TM - Thematic Mapper

Training - Informing the computer system which site to analyze for spectral properties or signatures; also called signature extraction.

Training Sites - Recognizable areas on an image with distinct (spectral) properties useful for identifying other similar areas.

U-2 - High altitude (remote sensing) aircraft

UHF - Ultra High Frequency

UNESCO - United Nations Educational, Scientific and Cultural Organization

Unsupervised Classification - Computer classification of digital images by placing similar pixels into categories without the aid of training-site data.

USLE - Universal Soil Loss Equation

UTM - Universal Transverse Mercator (Grid)

UV - Ultraviolet

Vector - A quantity possessing both numerical (scalar) value and direction.

VHF - Very High Frequency

VHRR - Very High Resolution Radiometer

VICAR - Video Information Communication and Retrieval (JPL software)

WBTR - Wide-Band Tape Recorder

WEFAX - Weather Facsimile

ZTS - Zoom Transfer Scope

Budget

PROPOSED BUDGET

P.I.s: J.E. ESTES
AGENCY: NASA, Supplement to NAGW-455
PERIOD: May 1, 1987 to April 30, 1988
TITLE: NASA Remote Sensing Information Sciences Research - Code EE Tasks

SALARIES*

J.E. Estes, Principal Investigator
    1 smr mo @ 100% @ 6738/mo   6,738

T.R. Smith, Professor
    1 smr mo @ 100% @ 6872/mo   6,872

J. Star, Assoc. Dev. Engineer
    2 mos @ 47% @ 3536/mo   3,324
    6 mos @ 50% @ 3713/mo   11,139
    4 mos @ 40% @ 3899/mo   6,238

TBD, Research Assistant
    2 mos @ 50% @ 1734/mo   1,734
    10 mos @ 50% @ 1821/mo   9,105

TBD, Staff Research Assoc IV
    2 mos @ 40% @ 2573/mo   2,058
    10 mos @ 40% @ 2702/mo   10,808

TBD, Senior Typist Clerk
    2 mos @ 33% @ 1367/mo   902
    10 mos @ 33% @ 1435/mo   4,736

SALARIES TOTAL: 63,654

BENEFITS

Estes
    6738 @ .1199 (smr)   808

Smith
    6872 @ .1914 (smr)   1,315

Star
    14,463 @ .29 (1987)   4,194
    6,238 @ .30 (1988)   1,871

Reseach Assistant
    2732 @ .0169 (smr)   46
    8108 @ .0135 (yr)   109

Staff Res. Assoc.IV

C-2

```
      8,543 @ .29 (1987)                          2,477
      4323  @ .30 (1988)                          1,297

Senior Typist Clerk
      3743 @ .29 (1987)                           1,085
      1894 @ .30 (1987)                             568
                                                --------
BENEFITS TOTAL:                                              13,770



EQUIPMENT
---------

Partial funding for purchase of
      image processing system                    15,000
      (including tax)
                                                -------
EQUIPMENT TOTAL:                                             15,000



TRAVEL
-------
One trip for 2 people for 3 days.  Santa Barbara
      to NASA Headquarters and Goddard Space Flight
      Center. (Washington, D.C.)

      1 RT airfare for 2 people @ $ 980 ea.      1,960
      3 days per diem @ $100/day for 2 people      600
      Car rental and transportation               100
                                                -------
                                                 2,660
One trip for 2 people to national meeting

      1 RT airfare for 2 people                  1,500
      3 days per diem @ $100/day for 2 people      600
      Car rental and transportation               100
      RT Car rental to NOSC                     -------
                                                 2,200

One trip for 3 people for 2 days. Santa
      Barbara to NASA/Ames Research Center,
      San Jose, CA.

      1 RT airfare for 3 people                    600
      2 days per diem @ $66/day for 3 people       396
      Car rental and transportation               100
                                                -------
                                                 1,096
TRAVEL TOTAL:                                                5,956



OTHER DIRECT COSTS**
--------------------
Materials and supplies                             500
Report preparation/publication                     418
Computer services:
```

```
    ITEL AS/6, VAX 11/780, VAX 11/750, and
       Broadband access fees.                          2,850
Telephone tolls & equipment                            1,550
Xerox and library reprographics                          500
Photographic and cartographic supplies                   900
Landsat Tapes from EROS & EOSAT                        3,000

                                                       -------

OTHER COSTS TOTAL:                                               9,718



        TOTAL DIRECT COSTS:                                   108,098

     ***INDIRECT COSTS @ 45% of MTDC:                          41,894
        (on 93,098)
                                                              -----------
        T O T A L:                                            149,992
                                                              ===========
```

   *Academic salaries are projected to increase by 5.0% per year for the
     period 7/1/87 and beyond. Staff salaries are also projected to
     increase by 5.0% per year for the period 7/1/87 and beyond.

  **Due to an internal recharge system, monies indicated under the salary
     section may be allocated and spent from "Other Direct Costs" as
     required by the system.

 ***This is the DHHS negotiated predetermined on-campus overhead rate for
     the period 7/1/85 to 6/30/90.

Summary

John E. Estes and Jeffrey L. Star

SUMMARY

John E. Estes and Jeffrey L. Star

This document represents a progress report of work conducted under grant NASA NAGW-455 during the period May 1, 1986 to January, 1987. This document also briefly describes the directions we propose to undertake in the coming year of this fundamental and applied research effort.

The Information System Research Group research continues to focus on improving the type, quantity, and quality of information which can be derived from remotely sensed data. As we move into the coming year of our research, we will continue to focus on information science research issues related to the Earth Observing System (EOS) of the Space Station Complex. The community of EOS users in the area of Earth Systems Science and applications represents scientists and researchers from many disciplines working on a variety of hardware and software systems at a number of geographically distributed locations across the United States and around the globe. Past research conducted under this grant has been used to extend and expand existing remote sensing research activities at UCSB in the areas of georeferenced information systems, machine assisted information extraction from image data, artificial intelligence, and vegetation analysis and modeling.

The program of research, documented in this progress report, is being carried forward by personnel of the University of California, Santa Barbara. Through this work, we have targeted

SUMMARY

fundamental research aimed at improving our basic understanding of the role of information systems technologies and artificial intelligence techniques in the integration, manipulation and analysis of remotely sensed data for global scale studies. This coordinated research program is possible at UCSB due to a unique combination of researchers with experience in all these areas.

Several of our projects have used this grant as a catalyst to aid other NASA offices in the research, in the integration of remotely sensed and other data into an information sciences framework. During this year we have received additional funds from NASA Code EI to supplement ISRG activities. In addition, we have conducted research for the United States Geological Survey on related matters and have received funds from NASA Code E to look questions revolving around Browse of large spatial data sets in the EOS Era.

Grant activities continue to generate papers in reviewed journals and to support graduate student Master's theses and doctoral dissertation research. We feel these activities are very important and we will continue to support student researchers and to submit material for publication in reviewed journals and national and international symposia. In addition, Grant researchers have been, and continue to be, asked to speak on EOS and Space Station related information system research across the country and abroad.

It is our belief that the work conducted under this Grant is

SUMMARY

significant. We are making progress as we prepare to take
advantage of the tremendous potential offered by the Space
Station Complex in general and the EOS system in particular.
Research such as this and that being supported by NASA at other
universities and NASA centers which make multidisciplinary Earth
System Science a reality and further the applications of the
earth observing satellite remote sensor systems of the 1990's.

Preliminary Design of a Farm Monitoring
Geographic Information System

David M. Stoms

Preliminary Design of a Farm Monitoring

Geographic Information System

The following chapter is a portion of the M.A. Thesis of Mr.
David Stoms. This work was funded in part by the current year's
funding, and represents a proposed design for an information
system. This geographic information system will integrate
remotely sensed imagery, field observations, and the local
farmer's detailed knowledge, as well as concepts from the field
of artificial intelligence.

Mr. Stoms continues in his graduate work in our laboratory
at this time, pursuing the Ph.D. degree. He plans to continue
his efforts on the integration of geographic information systems,
remote sensing technology, and artificial intelligence.

UNIVERSITY OF CALIFORNIA
Santa Barbara

Preliminary Design of
a Farm Monitoring
Geographic Information System

A thesis submitted in partial satisfaction
of the requirements for the degree of

Master of Arts

in

Geography

by

David Michael Stoms

Committee in charge:

Professor John E. Estes, Chairperson

Professor Frank W. Davis

Professor Julia Allen Jones

October 1986

ABSTRACT

Preliminary Design of
a Farm Monitoring
Geographic Information System

by

David Michael Stoms

The objective of this thesis is to improve our understanding of the need for and characteristics of an image processing/geographic information system (GIS) for monitoring crops on individual farms. A GIS is proposed that integrates remotely sensed imagery, field sensor data, and farmers' intimate knowledge of their operations to provide frequent, rapid information. Expert systems would perform some of the image analysis tasks, identify deviations from desired conditions, diagnose probable causes of stress, and recommend appropriate management responses.

For effective monitoring, remotely sensed inputs must be frequent, timely, of high spatial resolution, and with spectral coverage in the visible, near IR, mid IR, and thermal IR regions. Four satellite sensor systems (Landsat TM, SPOT, AGSAT, and the Earth Observing System) are identified as candidates to provide multispectral data. None, however, meets all of the specified requirements. Recommendations are made for modifications to each sensor system to make them suitable for farm monitoring applications.

This research outlines how, by putting remotely sensed data in an information system context, it can be made more useful for end-users not trained in image processing and interpretation. Development of this capability would be an important step in the dissemi-

nation of remote sensing and GIS technology to a broader range of users.

TABLE OF CONTENTS

viii

1. INTRODUCTION

The objective of this thesis is to improve our understanding of the need for and characteristics of an image processing/geographic information system (GIS) for monitoring crops on individual farms. A GIS is proposed that integrates remotely sensed imagery, field sensor data, and farmers' intimate knowledge of their operations to provide frequent, rapid information. Monitoring, as used in this thesis, refers to quantitative and qualitative assessment of general crop conditions, crop stage and biomass, plant stress, and expected yield. Expert systems would perform some of the image analysis tasks, identify deviations from desired conditions, diagnose probable causes of stress, and recommend appropriate management responses. This research outlines how, by putting remotely sensed data in an information system context, it can be made more useful for end-users not trained in image processing and interpretation. Development of this capability would be an important step in the dissemination of remote sensing and GIS technology to a broader range of users.

In the early days of digital civilian remote sensing, it was predicted that farmers would be major beneficiaries of this new technology (National Research Council, 1970). Data would be available to assist in making farm management decisions concerning such activities as irrigation, fertilization, and pest control. Since then, literally hundreds of articles have been published about large area applications of remote sensing in agriculture (summarized in Bauer, 1985; McDonald, 1984). However, remote sensing technology has not yet been applied to the daily management of individual farms. In fact, a recent article which details the history of agricultural remote sensing does not even mention farmers (McDonald, 1984).

A number of recent trends in remote sensing hold promise for making commercial

farm monitoring applications more feasible (Brammer, 1984). Among these advances are the development of:

-powerful, low-cost personal computers and display devices

-geographic information systems

-expert systems in remote sensing, image processing, and crop modeling

-sensors with improved spatial and spectral resolutions, and

-advanced data communications technology

Taken in combination, these trends create the potential for a system where individual farm managers can do their own automated image analysis on site, without need for centralized data processing and analysis. Further supporting this trend is the current emphasis by the Federal government on the commercialization of space, and a renewed interest within NASA in applications oriented research. In a recent proposal submitted to NASA for a Center for the Commercial Development of Space, commercial real-time farm monitoring was one of the primary research areas (The Ohio State University, 1986).

To date, there are no operational remote sensing systems that can meet farmers' requirements for frequent, real-time crop information. There certainly are no systems available that are simple to use by untrained interpreters. In fact, the features comprising a farm monitoring remote sensing system have not been clearly defined (Jackson, 1984).

A geographic information system is typically defined as an information system containing data referenced by spatial location (Estes, 1986). Inputs are usually in map or image form, but may also be attributes in tabular form. The first operating system was the Canada Geographic Information System (CGIS), implemented in 1964. It's original purpose was to locate marginal agricultural lands (Tomlinson, 1982). Since that time, dozens of systems have been developed, but only about ten are commercially available (Estes, 1986).

The commercial systems tend to be multipurpose and are not custom-tailored to farm monitoring.

Two aspects of a farm management information system are addressed in this thesis:

1) the elements of a farm monitoring image processing/geographic information system, including the role of remote sensing

2) the capabilities and limitations, within this GIS context, of four candidate satellite sensor systems for providing timely crop information

These two aspects are discussed in Chapters 5 and 6, respectively. The development of a prototype GIS, complete with the appropriate agriculture-oriented knowledge-based systems, is beyond the scope of this masters thesis research. However, several examples of monitoring cotton fields are described in Chapter 7.

To lay the foundation for the analysis, the literature is summarized regarding farmers' need for current crop information (Chapter 2), previous research into agricultural remote sensing (Chapter 3), and four candidate remote sensing systems (Chapter 4). The four satellite systems considered in this thesis are:

1) TM — the Landsat Thematic Mapper

2) SPOT — the French-built Systeme Probatoire d'Observation de la Terre

3) AGSAT — a dedicated agricultural system proposed by a group of commercial vendors and professors from Stanford University, and

4) EOS — the National Aeronautics and Space Administration's (NASA) Earth Observing System consisting of several instruments proposed for a Polar Platform on the Space Station complex

These four were selected because they are representative of existing and proposed systems that may be suitable for providing farm monitoring data.

It is assumed in this thesis that a farm monitoring system has practical commercial application as a management decision tool. Farmer interest in and cost-effectiveness of such

- 4 -

a system is not evaluated. Furthermore, it is assumed that most farmers are not trained in image processing and analysis, which implies that image processing tasks need to be automated. I also assume that the remote sensor systems being evaluated will operate as they currently exist or as they are proposed. This assumption covers not only the sensor package, satellite platform, and spatial, spectral and temporal resolutions, but the ground processing facilities as well. However, recommendations will be made in Chapter 8 on modifications to the sensor systems to make them suitable for farm monitoring applications.

Support for International Science

John E. Estes and Jeffrey L. Star

Support for International Science

This appendix is the text from a proposal to NASA Headquarters, which was submitted during this contract year. The purpose of this proposal is to examine the opportunities for collaboration and training in the international community. As a focus, we propose to work with the United Nations Environment Program offices in Nairobi, Kenya and in Geneva, Switzerland, with collaborators at NASA's Earth Resources Laboratory at NSTL.

Remote Sensing Research Unit
Department of Geography
University of California
Santa Barbara, California 93106

24 November 1986

Mr. Alexander Tuyahov
Code EED
NASA Headquarters
Washington, D.C. 20546

Dear Mr. Tuyahov:

Please consider this document an unsolicited letter proposal from the University of California, Santa Barbara, for work in support of international information science applications of satellite remote sensing technology. The Information Science Research Group (ISRG) at the University of California, Santa Barbara (UCSB) is currently conducting advanced research in the information sciences area for the National Aeronautics and Space Administration under two Grants (NAGW-455 and NAGW-987). Under these grants UCSB researchers are examining a variety of image processing database, networking, and data systems issues which will impact the earth sciences in the Space Station era.

A major focus of this research is facilitating access to a wide variety of regional, continental, and global scale Earth Science related data sets. This focus permits us to prepare for the analysis of imagery and other data from the Earth Observation System planned for the polar platform component of the Space Station Complex. While considerable attention is currently being placed upon access to U.S. national datasets by a number of NASA centers and NASA funded university research teams, it is our feeling that more attention to establishing ties to international data depositories is needed. As such this document represents an unsolicited proposal to examine the problems and potentials inherent in linking U.S. institutions to international organizations which process, maintain, and disseminate international and global Earth Science data sets.

Specifically, we at UCSB propose to work jointly with personnel of NASA's Earth Resources Laboratory, as well as the United Nations Environment Programs Global Resource Information (UNEP/GRID) program in both Nairobi, Kenya and Geneva, Switzerland to examine the technical issues and problems of hardware, software, data standards, and processing problems inherent in tying U.S. and international foreign data depositories into a functional system which can facilitate earth

scientists at UCSB and other NASA affiliated research institutions, to better access the data resident within these organizations.

UCSB researchers already have experience working with UNEP/GRID personnel. Under existing NASA grants at UCSB we have visited and obtained briefings on the operational status and future plans of the UNEP/GRID program in Geneva, Switzerland. UCSB personnel are also funded under NASA Grant NAGW-987 to visit Nairobi where briefings on UNEP/GRID will be held. In our contacts with Geneva UNEP/GRID personnel, they have discussed an interest in cooperating with ISRG/UCSB personnel in this effort.

One specific issue we would like to address is a need to augment the image processing and database management and manipulation capabilities at UNEP/GRID Nairobi. This will allow efficient manipulation and and transfer of data from Africa to UCSB. We believe that this will be necessary before meaningful analysis of the problems with and potential for international links with U.S. research organizations can be fully realized. In this proposed effort UCSB personnel propose to coordinate with NASA National Space Technology Laboratories (NSTL) personnel to conduct a preliminary analysis of the prospects and problems involved in linking UNEP/GRID Nairobi and Geneva into U.S. NASA related Earth Science research institutions. This international scale of research, involving elements of data set access, data standards, networking protocols, and so forth provide a good model for the information systems now being developed for the EOS era.

The Global Resource Information Database (GRID) is a new initiative of the United Nations Environment Program (UNEP) and an element of the Global Environmental Monitoring System (GEMS). Currently finishing its pilot phase, GRID grew out of a need to coordinate within a common geographic reference systems, the many environmental datasets that have been and continue to be collected by GEMS other UNEP programs and United Nations (UN) specialized agencies. GRID is being designed and implemented to facilitate access to and analysis of these data sets for environmental and resource assessments. As currently structured, GRID has two computer processing centers under UNEP direction: GRID-control at UNEP Headquarters in Nairobi, Kenya; and GRID-processor in Geneva, Switzerland. At present, the NASA Earth Resources Laboratory in Mississippi is the principal NASA center providing technical support and engineering expertise to GRID.

GRID-processor in Geneva is located in a University of Geneva facility provided to UNEP at no cost. It was designed and constructed specifically for GRID and the operation and maintenance of the facility is contributed by the Center of Geneva. The facility is well equipped with Prime, Perkin Elmer,

and ERDAS based data systems and image analysis processing capabilities. GRID-control, Nairobi has been established within GEMS offices in UNEP's Headquarters at the UN, Gigiri Complex. GRID-control functions within the administrative context of Earthwatch.

As previously stated, GRID is currently in its pilot phase. The pilot phase of the GRID program has four primary objectives. These objectives are:

1) to develop geographic information system (GIS) methodologies and procedures for constructing and manipulating global environmental data sets for the purpose of conducting environmental assessments. (The continent of Africa has been selected as the case study for this objective);

2) to demonstrate that GIS technology as applied within GRID is an effective tool which combines global and national data sets for resource management and planning applications at the national level. (The candidate countries for these demonstrations are China, Ethiopia, Indonesia, Kenya, Panama, Peru, Senegal, Sudan, Thailand and Uganda);

3) to establish the framework for cooperation and data exchange within international and intergovernmental organizations which deal with environment-related matters, such as FAO, WHO, WMO, ICSU, ILCA, IUCN, etc.; and,

4) to provide training opportunities in GIS and resource data management technologies employed by GRID to the scientists and resource managers from participating developing countries.

The pilot phase of GRID originated as a cooperative project between the GEMS Programme Activity Centre of UNEP and the National Aeronautics and Space Administration (NASA). It grew out of common interest of the two organizations. UNEP is mandated to monitor and assess the state of the global environment on a continuing basis; NASA is one of the principle organizations involved in studying the long-term changes in processes that affect life on the Earth. From this beginning other nations and organizations have been brought into the venture so that GRID has become a truly international cooperative effort. It is our feeling that it is now time to include the NASA related university community in this effort. The research proposed herein represents a preliminary step in this direction.

UCSB has been involved in a number of efforts over the past years which directly impact the requirements and opportunities

for International dataset production and exploitation. In May
1985, NASA's Office of University Affairs signed a grant
establishing a Remote Sensing Information Sciences Research Group
at UCSB. Research conducted under this ongoing grant has been
used to extent and expand UCSB's efforts in the areas of
georeferenced information systems, machine-assisted extraction of
information from image data, artificial intelligence, and
modeling. The group's philosophical emphasis is to improve the
type, quantity, and quality of information which can be derived
from remotely sensed data.

Two examples of our efforts under this grant have direct
relevance to the problems of global dataset creation and
dissemination. In April 1985, we completed a brief study which
considered the applications of artificial intelligence techniques
to large distributed networks. We identified a number of areas
where such technology could provide more efficient problem
solving, communications, and data base creation and sharing than
conventional techniques (Dubayah, Smith, and Star, 1985).

Another example of advanced technical applications involves
a fundamentally new geographic information system, developed
under shared funding from NASA, United States Geologic Survey,
and Digital Equipment Corp. This new system, KBGIS II,
incorporates new developments in data structures, software
engineering, and artificial intelligence. The system provides a
user-configurable applications interface, one-step search for
complex objects, and state-of-the-art capabilities for
intelligent search of exceedingly large datasets. Such a system
would be a natural environment for the manipulation and analysis
of global-scale data.

UCSB has just received funding from NASA Headquarters for a
new effort which also has direct relevance to the development and
use of global datasets. This new grant (NASA NAGW-987, "Browse
In the EOS Era") permits to work on the key issues of browse of
remotely sensed and other ancillary data. Earth science and
relevant technology have developed to a point where a number of
agencies, including UNEP, are actively considering the
development of large area datasets. These large, heterogeneous
spatial data holdings require new insights in database management
and processing, to be able to serve a wide range of
geographically distributed users. Our work under this new grant
will include developing a set of functions that permit a user to
examine remote databases, to be able to make an intelligent
decision on when these databases contain useful material for a
variety of different processing and analysis functions.

There is an international need to improve access to global
data sets for NASA and NASA-related researchers both in the
United States and abroad. The UNEP/GRID activity is a source of

such data, of growing significance to both research and operational agencies. It is important that we begin to facilitate interactions to assess the problems and prospects for connecting into this important source of global datasets. The material which follows presents our methodology and research requirements for beginning this important task.

METHODOLOGY

1   Analysis of existing systems in Geneva and Nairobi.

     In our contacts with staff at GRID's Geneva facility and NSIL personnel, under NASA Grant NAGW-455, we have examined and discussed the primary functions required at UNEP/GRID Nairobi. We will determine whether sufficient hardware is available in Nairobi, and advise them on necessary staff development.

2   Determination of requirements for system augmentation at Geneva, Nairobi, and Santa Barbara.

     One of the principal constraints of new systems at these location is compatibility with existing hardware and software at principal NASA and UNEP/GRID facilities. We will continue our analysis jointly with staff at NSTL, NASA Headquarters, and UNEP/GRID Geneva.

3   Subject to funding, purchase required hardware and software.

4   Install and integrate new capabilities

     Our collaborators at NSTL have detailed experience in the specific hardware and software environment installed at GRID Geneva. Their specific participation in this proposed effort will include acceptance tests of the equipment at the shipping location. NSTL has also agreed to ship the equipment to Nairobi, as well as provide staff time and travel for installation and maintenance during the contract year.

5   Conduct pilot tests of data transfer to document the capabilities of the new components.

     UCSB personnel will travel to Nairobi to begin staff development and training in both image processing and geographic information systems technology. We will also provide training material in these areas, for use at UNEP/GRID Nairobi. This training material will include both reading materials and sample datasets. The sample datasets will be chosen after discussions with our collaborators at NSTL and UNEP/GRID, and can form a basis for future staff development and training courses.

6   Prepare final report and submit to NASA Headquarters.

Based on these meetings, as well as discussions with our other collaborators and technical monitors, we will begin preliminary analysis of the problems and prospects involved in efficiently linking these different organizations. This analysis will result in a report, detailing the analysis and the suggested alternatives for efficiently moving data between the laboratories at NSTL, Geneva, Nairobi, and Santa Barbara.
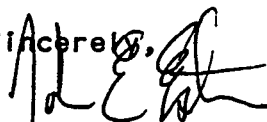
DELIVERABLES

UCSB will purchase an appropriate computer system, based on requirements and constraints identified by ourselves and collaborators at NSTL, NASA Headquarters, and UNEP/GRID. Shipment and installation of this system will be provided by NASA's Earth Resources Laboratory, NSTL. Training in the operation and applications of this system will be undertaken jointly by NASA and UCSB staff.

UCSB will provide a final report of our activities with respect to this grant. This will detail our meetings during the project, our analysis of the utility of the computer in Nairobi, and tentative plans for future collaboration and support activities for the effective dissemination of NASA-developed technology to the developing countries now being assisted by the UNEP/GRID program.

We look forward to the opportunites which this grant can provide, if funding becomes available.

Sincerely,

Prof. John E. Estes
Department of Geography

Browse In the EOS Era

John E. Estes and Jeffrey L. Star

Browse in the EOS Era

This appendix is the text from a proposal to NASA Headquarters, which was submitted during this contract year. The first year of this proposal has been funded, and we have just started what we hope to be a three to tour year ettort. Our work under this new effort for NASA interfaces well with our Office of University Affairs grant. The Browse program is designed to provide us with experience in problems which will become more acute in tne timetrame of the EOS platform. Under tne Browse program, we will develop a testbed, in which we will examine some of the problems of geographically distributed archives, and tne problem of determining the scientific value of remotely sensed data located at a remote archive.

ABSTRACT

The Information Science Research Group, University of California, Santa Barbara, proposes to supplement their existing National Aeronautics and Space Administration grant, number NASA NAGW-455, to conduct basic and applied research on browse in the EOS era. The newest and most important initiatives in the U.S. civilian space program revolve around the Space Station Complex. From the perspective of scientists studying the dynamic coupling of the lithosphere, biosphere, hydrosphere, and atmosphere, the most important component of the Space Station Complex is the Earth Observing System (EOS).

Data volumes projected for EOS are extremely high. The high data volumes, coupled with the geographically distributed nature of the EOS user community, calls for innovative research on access to EOS data. It is most important to be able to browse EOS datasets and select only that portion of the data relevant to a given research problem.

This unsolicited proposal develops a three-year phased research program. Our program is aimed at the development of a robust, well-structured system for browse, able to support users of different levels of sophistication.

There are two primary functions imbedded in the testbed we propose to develop: locating relevant data, and viewing the data. We will emphasize the latter in our proposed efforts. Locating data involves four general types of information about the data: spatial location, information about the data object, themes of coverage, and information about access to the data itself. We will rely in part on ongoing efforts in this general area, including those at NSSDC and the NASA Pilot Data Systems. The second function, viewing the data, includes at least three levels of access: viewing the data itself, viewing attributes of the data, and viewing information derived from the data.

During the first year of this proposed effort, we will hold meetings of selected scientific users, conduct function tests and system development, and we will begin collaborations with other data archive sites. Results of these activities will be documented in progress and annual/final reports, and in the reviewed literature.

In conclusion, the concept of EOS as an information system is an important philosophical direction for the program. EOS will significantly advance our understanding of our global system, as data is provided to scientists in many disciplines around the world. Browse of datasets is a key element in facilitating scientific access to EOS data.

INTRODUCTION

This document represents an unsolicited proposal for a supplement to National Aeronautics and Space Administration (NASA) Grant Number NASA NAGW-455 to conduct basic and applied research on key issues related to browse of EOS and related ancillary data by interdisciplinary science users. In May, 1983, the office of University Affairs of the National Aeronautics and Space Administration (NASA) signed a grant establishing a Remote Sensing Information Science Research Group at the University of California, Santa Barbara (UCSB). Research conducted under this grant continues today and is aimed at improving the type, quantity and quality of information which can be derived from remotely sensed data. During the past year, the basic and applied research conducted under the auspices of this grant has begun to focus on the information science research need of the Earth Observing System.

As the world's population expands, there is an ever increasing need for systems and techniques capable of acquiring, integrating, communicating and analyzing information concerning the extent, use of and change in major components of the earth's surface. NASA is playing a key role in the development of such systems and capabilities. This role requires that forsighted fundamental and applied research be directed towards the scientific application of technologies which can improve the base upon which assessments may be made of both the current and changing status of the components of the biosphere, hydroshere, lithosphere and atmosphere.

The proposed Earth Observing System (EOS) can play an important role in improving our understanding of the spatial and temporal dynamics of major components of our global life support system. EOS, as defined by the Science and Mission Requirements Working Group (NASA, 1984a) consists of: "a suite of instruments in low Earth orbit acquiring measurements of the Earth's atmosphere, surface, and interior, an information system to support scientific research, and a vigorous program of scientific research stressing study of global-scale processes that shape and influence the Earth as a system". EOS is being designed to produce large volumes of multispectral, multitemporal data for multi-disciplinary investigations. An important consideration of the EOS program then, is the development of an infrastructure and mechanisms to speed delivery of data from EOS to the science community. A key aspect of the capability is the development of a mechanism to insure that investigators are not deluged with unnecessary or redundant data. A mechanism which can aid in minimizing the potential of this happening is to develop ways and means to permit scientists to browse EOS data and select only those data most appropriate to their research.

The unsolicited proposal which follows describes a three year phased research effort to address the question of browse in the EOS time frame. Although a three year effort is proposed, we are at

present requesting supplement of our existing NASA grant (NASA NAGW - 455) for the first year of this effort. Funding for the second and third years will be requested later, and we understand that such funding is contingent upon both the performance of the research and funding availability.

After this introduction, a background section describes the need for the EOS system and its importance to global science. This is followed by a conceptual framework for the proposed research on the question of browse. This section includes approaches to the study and descriptions of our proposed three year-phased research effort. A conclusion section is next, followed by the budget for this effort. Vita of key project personnel and descriptions of facilities and equipment at UCSB are also included. An appendix, describing a related effort, is attached at the end of this document.

It should be emphasized that this research will take advantage of the already extensive interdisciplinary contacts and research on Space Station and EOS related activities currently being conducted at UCSB. UCSB researchers are working directly in the area of land, climate and oceanographic research. In addition, we are cooperating with and/or working on research for the U.S. Geological Survey and the Research Libraries Group (see appendix) on topics related to improving the scientific utility of spatial data. Finally, the research effort proposed herein will gain maximum benefit from membership on important data system activities committees and working groups on the part of project scientists. These committees and working groups include: the Committee on Data Management and Computation; the Space Station Data User Working Group; EOS Data Systems Panel; Pilot Land Data System Science Steering Group; Research Library Group Task Force on Geo-Information; and the Task Force on Scientific Uses of Space Station Earth Science Panel. Membership and participation in these efforts will continue to provide insight as the research we propose moves forward.

BACKGROUND

Earth science and associated technology development have progressed to a point where the conduct of global science appears feasible (Estes and Star, 1986). This scientific and technological development has been caused, in part, by the emergence of problems which require a multidisciplinary approach (NASA, 1984a). The need to reach an improved understanding of problems such as as deforestation, dessertification, changes in atmospheric carbon dioxide, and acid deposition (to name a few) grows more important with each passing year. The understanding of these problems, however, requires new approaches to science, which combine multidisciplinary teams employing advanced technologies which can generate a type, quantity and quality of data not previously available to the scientific community. The proposed Earth Observing System (EOS) currently being planned for the polar-orbiting platform of the United States Space Station complex offers

this potential.

The newest and most important initiatives in the U.S. civilian space progam currently revolve around the Space Station Complex. The Space Station complex includes the core station, its associated co- and polar orbiting platforms and information systems. This proposed suite of platforms and support systems offers a unique potential for facilitating long term scientific investigations on a truly global scale.

Basically, the man-tended systems which are proposed for the various Space Station Complex platforms have the capability of providing a wide range of data from both operational and research sensors. These large volumes of multispectral, multitemporal data, which will be generated by the proposed systems, must be supported by efficient and effective information systems to provide the potential for data continuity which has, to a large degree, been lacking from sensor systems ooperating on independent free flying platforms. The challenge to the remote sensing community is, in essence, two-fold. The first challenge is to get ready to handle the large volumes of data which will become available in the 1990 time frame. The second challenge is to bring the science and technology we are developing to broader constituency, in the service of global science (Estes and Star, 1986).

From the perspective of scientists studying the earth's surface and atmosphere, the most important component of the Space Station complex is the Earth Observing System (EOS) (NASA, 1984a; NASA, 1984b). EOS, based on the current design concept, has both active and passive earth surface imaging systems as well as atmospheric sounding systems. EOS represents an evolutionary step in our capabilities for remote sensing of the earth. When implemented, EOS may provide the earth, ocean, and atmospheric science communities with data to support unprecedented integrated investigations among disciplines and scientists from many nations. Unlike the previous generation of satellites, designed largely for relatively limited constituencies (e.g., Landsat for the land scientist, Seasat for the oceanographic community), EOS has the potential to provide an integrated source of information which recognizes the problems and issues which involve the dynamic coupling between the oceans, land surface and the atmosphere.

In the same way that EOS represents an evolution in earth viewing satellite technology, we believe that the scientific objectives which EOS may help to advance can produce an evolutionary improvement in our understanding of our planet. Traditional branches of the earth sciences have, in the past, generally focused on problems which are limited in scope to modest areas and to relatively narrow ranges of biophysical, geochemical and socioeconomic processes. Traditional science has been limited by the capabilities of existing technologies to measure, map, monitor, and model key elements of these processes. EOS datasets will be able to expand our capabilities in these areas,

and will facilitate new scientific insights concerning the world we live in.

EOS has the potential to foster expanded collaboration within the scientific community towards interdisciplinary science on an international scale. Yet, if we are to fully employ the potential of EOS it must be done within an information systems context, linking scientists together with both required facilities and each other. Such an approach can improve the global science community's access both to data sources and processing capabilities. Global science is a data-intensive activity and in its broadest sense, EOS, as an information system, can provide a tool for improved understanding of our planet (NASA, 1984a). There are, however, a number of unanswered questions concerning the operational and commercial uses of the sensor systems on polar platforms.

Central to these questions is science access to the data from the EOS systems. Indeed a key recommendation of the EOS Science Mission Requirements Working Group is the development, as soon as possible, of "a data system that provides easy, integrated, and complete access to past, prestent and future data". Data volumes projected for the EOS system are high. EOS Synthetic Aperature Radar (SAR) and High Resolution Imaging Spectrometer (HIRIS) data rates are projected between 700 and 800 megabits per second. Daily data rates of one terabyte per day have been suggested. To illustrate this size more dramatically, one terabyte is a shelf of floppy disks roughly seven and one-half miles long, or a 400 square foot room of 1600 BPI tapes.

The real challenge to the international science community in general, and NASA in particular, is the development of the infrastructure to facilitate the operation of the EOS Information System. Scientists must be able to achieve rapid access to needed data. The ability to browse data and efficiently and effectively pre-judge its potential utility for a given investigation is a critical need. The material which follows, then, represents an unsolicited proposal for a three year phased applied research effort to study important issues related to browsing remotely sensed ancillary data sets of interest to future EOS science users.

CONCEPTUAL FRAMEWORK.

The stated goal of the EOS Data and Information system, is "to meet the challenges of EOS mission operations, data transport, processing and ata management, in addition to the challenges of access to information and data not under direct control of EOS or even NASA" (NASA, 1986). This requires the development of a geographically distributed information system capable of handling the large volumes of data which EOS will produce.

It is anticipated that by the 1990's, local processing capabilities combined with network technologies will allow such a spatially distributed system to become a reality. A key objective of the EOS Information System is to provide local and remote interactive access to a variety of capabilities, data sets, and resources. With data loads projected as high as one terabyte per day, the ability to browse datasets prior to retrieval must be an important function for the EOS Information system. Current estimates of the need for such a browse capability, of the type we propose to investigate, ranges between 1500 and 15,000 users (NASA 1986).

Because of the large data volume which can be generated by EOS, coupled with user requirements for ancillary data for test sites around the globe, user-friendly human interfaces to the information system are necessary. Users will require a variety of kinds of assistance to be able to find potentially relevant data resources, and evaluate their potential importance. Knowledgeable users may have detailed information about data sets, archives, and processing capabilities; these users may prefer a terse, streamlined interface to the system. Less knowledgeable and infrequent users of the system may require more detailed on-line user assistance, including help facilities and menu-driven interfaces. Such users may require detailed information about data set characteristics such as sensor parameters, and information such as spectral, temporal and geographic characteristics of available datasets – both for EOS-produced data and important ancillary data.

These requirements indicate the need for a robust, well-structured system, able to handle users of different levels of expertise about both EOS sensonrs and the data system itself, as well as users from a range of disciplines. The ability to browse data sets will be central to the needs of both advanced and less-sophisticated users.

BROWSE

There are two functions to examine when considering a system able to browse distributed information sources. First, a future EOS Information System must provide facilities to locate data which is potentially of interest. This includes both identifying archives or repositories where the data may be stored, as well as selecting the relevant items from within the archive based on appropriate geographic location, sensor characteristics, etc. Second, we must be able to view either the data itself, or relevant aspects of the data set. We will emphasize the latter function in our proposed testbed development.

Locating Data

Data archives and repositories throughout the scientific community are beginning to understand how to make data easily available to users. One of the first steps is to prepare data <u>directories</u>, which are

essentially lists of the available datasets which are kept in the archive. This is sometimes called metadata: data about the data. Such metadata is often kept in a data base management system (DBMS), to permit easy access to the contents of the database, as well as manipulation of the directory to find relevant material.

Important efforts at interfaces to large data collections are beginning through several agencies, including NSSDC, the NASA Data Pilot programs, USGS, EOSAI, SPOI, and the Research Libraries Group. In parallel with these development efforts, we are seeking provate foundation funding to automate a directory to the UCSB Map and Imagery Collection - a collection of approximately 3.2 million image and 300,000 maps. We propose to use the results of these new developments, collaborating with these research and development efforts where our staff expertise will be helpful.

There are four general types of information we would like to be able to consider when looking for a specific element of data. These are:

- locator:

geodetic coordinates that specify the area of interest. For locations on the earth's surface, a latitude/longitude pair is the most universally understood coordinate system. Auxiliary coordinate systems, such as the Landsat path/row system, should be included where appropriate. Minimum useful information would be the center of an object; more useful would be a series of (x,y) locations for the boundary.

- object:

properties of the data object itself. This includes a unique standard descriptor (i.e., scene ID, map name), date the data was compiled or acquired, and an indication of data quality.

- themes of coverage:

for a multispectral sensor, this might include the specifications of the spectral bands (i.e, bandwidth and center passband wavelength); for a derived product, this might be the algorithm, or the resulting classification (i.e, transformed vegetation index, or soil moisture).

- access:

the location of the data itself, as well as means to acquire or view the data. This would also include information about the data format.

An example of a directory to hold information about remotely-sensed data might look like the following:

Scene Center	Path/ Row	Sensor	Scene ID	Scale	Date	Cloud Cover	Access	Format
35 15'N 119 10'W	10/50	TM	840425768	30m	4Jan85	10%	EOSAI	1600BPI BSQ

This kind of organization of information about the data holdings
suggests a relational database structure. In particular, we should
emphasize that each field in this dataset could be the start of a
search for data. For example, we could begin by specifying the ground
location of a test site, or by specifying a range of dates, or a
combination of sensor and data quality. Full boolean operations on the
data fields must be possible.

There are systems that already contain this kind of metadata.
Commercial DBMS's may be a good start towards building a useful
capability for the scientific community. It is unreasonable to expect
that existing data collections will necessartily place their data
either within a future Earth Observing Information System (EOIS), or
obey some not-yet-determined standards for interface to the EOIS. Two
issues immediately confront a potential user: gaining access to
distributed directories, and working with these distributed directories
as it they were homogeneous.

For this proposed work, it is important to stay aware ot NASA's
near-term and medium-term networking plans (i.e., PSCN). We believe
that access to distributed directories and catalogs can be dealt with
at several levels. For scientists with frequent need to access the
systems, dedicated attachments to NASA's networks may be appropriate.
For those in the community with less-frequent needs, dial-up and batch
access facilities may meet their requirements inexpensively. We
propose to provide a free telephone number (800 or WATS service) to our
facilities for our collaborators in this project, to exercise this low-
cost option for providing service.

Any number of institutions are working towards on-line query of
catalogs and data directories (i.e., UCSB Map and Imagery Laboratory,
NSSDC, EOSAT, USGS, SPOI). It is not reasonable to expect that a given
user understand the details of working with a large number of separate
data directory systems. For spatial data, we believe that, while the
explicit syntax of data base queries will vary between systems, the
underlying semantics are quite restricted. This presents an
opportunity for a browse facility to appear to be homogeneous to an end
user, where in fact the databases are distributed around the country,
and there are different database query languages involved.

Approaches that should be studied include:

1 - developing a standard for data catalog and directory query

2 - developing software to convert data catalog and directory queries

to and from a specified set of query languages

3 - constructing a central catalog/directory node on a network, which routes queries to appropriate databases and translates between query languages

4 - developing transportable software to run on local processors for network routing and query translation

5 - automated assistance to efficiently decompose complex queries to minimize search and network costs

6 - automated assistance to make the system friendly and efficient to a wide range of users

Viewing Data

Once relevant data sets have been located through the mechanisms above, the user must have some capability to examine the data. Some simple forms of examination may be built into the directories described above. For example, data quality may be one of the fields in the data directory. A useful distinction can be made between data kept online for some period of time, in comparison to offline materials that would require some operator action before viewing is possible (i.e., mounting a tape or optical disk, duplicating a sheet of microfiche).

We can structure the viewing of data into three levels:

1 - viewing the data itself. This may involve making a portion of the data available to the user via on-line methods, or pre-computed browse products may be made routinely (such as Landsat microform browse images today).

2 - viewing data attributes. This may involve information in the data directory (as in the digital data quality), or possily in an image header (ephemeris data, for example).

3 - viewing derived data. This ranges from information derived from the data values (such as mean and variance of specified spectral bands), to more sophisticated derived products (such as data rectified to some map projection, or a transformation of the data according to some model). Useful derived data for one discipline may not be particularly useful to others; the meetings we propose in the first year will help us to identify the needs and desires of the NASA and NASA-related science communities in this area. While the needs of operational (e.g., NOAA) and commercial (e.g., EOSAT, SPOT) users are important, we feel that focusing on the needs of the diverse NASA science community will provide an appropriate test of the system for this proposed effort.

Once the user has identified the desired characteristics of the data to be viewed, there are a number of specific questions that arise.

University of California, Santa Barbara   Page 10

We propose to address several of these during the first year of this effort, in conjunction with the user's meetings we propose. These questions include:

1 - are there some set of viewable products that might either be pre-computed and stored, or where dedicated hardware might speed their extraction?

2 - for different disciplines, are there data transformations that might be applied to the viewable data that enhance the data's value?

3 - what level of data compression might be applied to the data, balancing data resolution, time, network costs, pre-processing costs, and local data reconstruction costs?

4 - are there relatively standard hardware configurations that the browse system must support at various levels of performance?

5 - are there specified data formats already available for the data that must be passed between systems?

6 - what levels of services can be provided to users with different levels of sophistication and data processing hardware?

For a first level analysis, we anticipate providing different levels of service based on at least 3 different kinds of hardware at a user site:

1 - dumb alphanumeric terminal
2 - smart alpha terminal with known video attributes and addressable cursor
3 - graphics workstation

1 - 'Dumb' terminal query of the database.

Minimum hardware - almost any dumb terminal, hardcopy or soft.

Command-driven SQL-like interactive session:

```
---------------------------------------------------
I FIND (IMAGE) WHERE (LOCATION = 30N 116W)  I
I                WHERE (SCALE > 1:24500)        I
---------------------------------------------------
```

2 - Terminal query with smart prompts.

Minimum hardware - alphanumeric terminal with cursor addressing

prompted query language:

```
----------------------------------------------------------
I FIND ( ANYTHING          WHERE ( LOCATION          I
I      ( MAP                     ( DATE              I
I      ( IMAGE                   ( SOURCE            I
I      ( DIGITAL DATA            ( DATA QUALITY      I
I                                ( SCALE             I
I                                ( THEME             I
I                                ( FORMAT            I
----------------------------------------------------------
```

The screen might present a menu of this form, and the user can move a
highlight around the screen to select items. When an item is
highlighted, a pull-down screen helps prompt for reasonable answers:

```
----------------------------------------------------------
I FIND ( ANYTHING          WHERE ( LOCATION          I
I      ( MAP                     ( DATE              I
I      ( IMAGE                   ( SOURCE            I
I      ( DIGITAL DATA            ( DATA QUALITY      I
I                         ==>>   ( SCALE      <<==I
I                                                   I
I                                =================  I
I                                I SCALE OPTIONS: I I
----------------------------------------------I  1:1 000 000    I--
                                              I  1:  250 000    I
                                              I  1:  100 000    I
                                              I  1:   62 500    I
                                              I  1:   24 500    I
                                              I please enter    I
                                              I your request;   I
                                              I < & > are       I
                                              I permitted.      I
                                              I _____   I
                                              =================
```

And of course, pull-down menus should be available at any time for
help. We emphasize that menus can be a help to a less-experienced
user, and a burden to an expert. User interfaces must provide both
friendly, english-language menu-level assistance to some users, and an
efficient, terse mode of access to others.

3 - Graphics terminal

    Here, the smart prompting may be combined with a base map display
and interactive graphics to select geographic regions of interest.
Local intelligence in the graphics terminal can help drive the system,
minimizing network traffic and response time.

    The base map display must be capable of zoom/roam, with real-time
reporting of the latitude/longitude of the cursor and any selected
region of interest. The user has the option of highlighting the
effective ground location of any object that is retrieved from the

database.

For example, using the smart menu above, the user could narrow the search to sensors with specfied spectral bands, as well as a specitied range of IFOV. Next, the user calls up a base map of the U.S., zooms into Central California, and paints a round area of interest by specifying the dimension of the region. The system derives the appropriate latitude/longitude window of interest from the displayed graphics, and starts the search. Centerpoints (or optionally, outer boundaries) of the retrieved records are plotted on the base map display. The user can then highlight the graphic representation of a particular object, and get a screen display that tells about that particular object.

Options to continue narrowing the search, change the parameters of interest, and let search parameters be based on other searches ("find me a map that corresponds to the last image you found for me") need to be included.

These views of the operation of an integrated browse system are only a framework. We believe that we must work in conjunction with users in ditterent disciplines, as well as in conjunction with those institutions developing data catalogs (for example, NSSDC and NOAA/NESDIS) and directories. In this way, we can adapt our development efforts to take advantage of efforts (such as the NASA Data Pilot development activities) that are leading towards EOS and the EOIS in the 1990's.

WORK PLAN

We view our efforts as a part of a three-year period of basic and applied research and development. While this propsal discusses our concept for the three year phased effort, we have proposed a budget for only the first year. We understand that tollow-on funding is contingent on both our performance and the availability of funding.

Year 1 provides us with necessary user input, while developing tools for the community at large. Year 2 begins the principal development efforts, as well as consideration of advanced technology approaches where appropriate. Year 3 includes implementation and testing of a pilot browse system for a designated group of users. We outline the three years in brief in the following paragraphs; our budget in this proposal covers only the first year. We anticipate that the second year will require somewhat higher levels of funding that the first year, with requested expenditures during the third year at first year levels.

<u>Year 1.</u>

   <u>User Meetings:</u> We propose to convene two meetings of selected scientific users (from NASA laboratories as well as the University of Michigan, University of Wisconsin, and Purdue University, among others) and staff from potential data set repositories and archives (JPL/Ocean Data System, NSSDC, PLDS, EOIS, NOAA). The first meeting will, in part, set an agenda for the second, larger meeting. In conjunction with representatives, this planning meeting will identify key issues and develop position papers which will form the basis for discussion at the later meeting.   The purposes of the larger meeting are to identify subsets of the needs of users in different disciplines, and identify the nature and range of services that may be available from agencies with major data holdings in the near-term as well as in the Space Station Era. We also propose to send one senior staff member to the International Symposium on Remote Sensing of the Environment in Nairobi. This conference is sponsored in part by the United Nations Environment Program, and an excellent opportunity to make contact with the International community, to acertain and understand their interests in browsing large geographic datasets.

   <u>Function Test and Development:</u> A variety of generic functions are necessary to a browse system. In the first year, we will begin to develop and test these functions, using colleagues both in NASA and with NASA-funded research interests as "clients". This involves developing a test bed configuration, accessible to outside users (via NASA networks if available, as well as dial-up 800 WATS service). Tests involve data compression, data display, graphics interface proceedures, and data format and communications standards.

   <u>Collaboration with Other Sites:</u> In part as a result of the user meeting, we hope to formalize collaboration with other users working in similar areas, as well as with staff in agencies with significant data holdings. These collaborations must include shared software development and testing, dataset exchange, and bench tests of developed user aids and software components.

   <u>Initiate Planning for Future Developments:</u> Towards the end of the first year, we will develop detailed plans for the second year's research efforts, outlined below. In addition, we will begin discussions with Headquarters and other NASA personnel on future project deliverables.

<u>Year 2.</u>

   <u>Principal Software Developments:</u> Software to be able to browse data catalogs and directories from remote sites. Software to be able to view data stored at remote sites. Software for data query language conversion.

University of California, Santa Barbara   Page 14

Exploration of Advanced Technology: Consideration, and where appropriate, development of specifications for advanced technology (such as artificial intelligence techniques) for the test bed. Functions that might benefit from this include data compression, query decomposition and translation, distributed problem solving, network operation and management, and data format conversion.

Users Meeting: Based on those collaborations begun in the first year, monitor the effectiveness of tools and techniques developed in the early phases of the program.

Testbed Evaluation: Provide access to specified distributed datasets from the testbed, as well as access to the testbed by specified users offsite. Evaluate the performance of the system, in terms of meeting user requirements in the near-term as well as in terms of developments in NASA's network facilities and in the EOS system itself.

Year 3.

In the final year of this proposed effort, we will continue the work begun in the first two years, and integrate advanced technology developments into the testbed. In particular, we will continue our examination of NASA Data Pilots and Information systems. We will make recommendations to NASA headquarters on integration of the results of our work into ongoing data pilot activities. Further, in this final year we will provide software, and assist software development for specified users and agencies providing data. Finally, we will make our testbed system available to our collaborators over a network, for final system test and evaluation.

PROPOSED PROJECT DELIVERABLES

There are several products to be provided as a part of this proposed research effort. A semiannual progress and annual or final report will describe the bulk of our efforts. The semiannual progress report will include discussions of work accomplished to date, summaries of user experiences with the testbed system, and discussions of work and a budget for the following year. Also included will be copies of reviewed journal articles coming from this research effort. The final report will include discussions of all the work accomplished during the period of the grant.

Additionally, copies of software and relevant documentation developed under this proposal will be made available to to the sponsor.

CONCLUSIONS

The challenge before NASA and the International scientific community is to continue to develop both the infrastructure and expertise which will insure that the EOS Information System works properly. On the one hand, we must continue to develop the science and technology of remote sensing. This includes research on improved communications and advanced processing techniques, to natural language interfaces and advanced scientific workstations, as well as new sensor technology. The proposed research effort on browse of EOS and ancillary data embodies, to some extent, all of these. In addition, the work to be performed in this proposed research effort provides a bridge to the science community. This bridge, if effectively built, will provide a pathway which can speed our progress towards global science and an improved understanding of the dynamic of our global life support systems. This improved understanding is the cornerstone, the true goal of the EOS system.