NASA-CR-173856
19840023669

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

Final Report Year 1

Dr. John E. Estes
Principal Investigator

Co-Investigator/Contributors:

Dr. Daniel Botkin
Dr. Donna Peuquet
Dr. Terrence Smith
Dr. J. L. Star

May 1, 1984

ISRG

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

Final Report Year 1

Dr. John E. Estes
Principal Investigator

Co-Investigator/Contributors:
Dr. Daniel Botkin
Dr. Donna Peuquet
Dr. Terrence Smith
Dr. J. L. Star

May 1, 1984

ISRG

N84-31739

Table of Contents

Introduction

Prof. John E. Estes and Dr. Jeffrey L. Star

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

In May 1983, the Office of University Affairs of the
National Aeronautics and Space Administration (NASA) signed a
grant establishing a Remote Sensing Information Sciences Research
Group at the University of California, Santa Barbara. This
document represents the first annual report of work conducted
under this grant (Grant# NASA NAGW-455).   This document also
briefly describes the research to be conducted in the second year
of this effort.

Fundamentally our research will continue to focus on
improving the type, quantity, and quality of information which
has to date and can be derived from remotely sensed data.   This
effort, as seen in the following text, has been directed at
integrating, extending, and expanding existing remote sensing
research activities at UCSB in the areas of Vegetation analysis
and Modelling, Georeferenced Information Systems, Machine
Assisted Information Extraction from Image Data, and Artificial
Intelligence.

As the world's population expands, there is an ever
increasing need for systems and processing 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
such capabilities.   This role, however, requires that farsighted
fundamental 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, hydrosphere, lithosphere, and

atmosphera.

The program of research, documented in this annual report which is being carried forward by personnel of the University of California, Santa Barbara, represents the start of what we consider to be a five to ten year effort. 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 as UCSB has a unique combination of researchers with experience in all these areas.

Efforts during the first year of this effort have 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 involved in the establishment of a Pilot Land Data System for NASA Headquarters.

In addition we have begun a small pilot project, building on our existing work with NASA Life Sciences and Earth Sciences and Applications Divisions through the Johnson Space Center. This pilot study is being used to test our concepts of how artificial intelligence assisted information processing, extraction, and user interface techniques may be integrated to increase our ability to conduct studies which lead to an improved understanding of global processes.

As will be seen in the detailed sections to follow,

considerable field work has been completed, in concert with

Johnson Spacecraft Center (JSC) staff, to gather the necessary

data. The development of a "test bed" processing environment

will begin in the coming year of the project, serving to both

demonstrate the potential of these technologies and to improve

our interactions with NASA Headquarters and Centers personnel.

We feel strongly that such close interactions are essential as we

begin our five to ten year research program.

In addition to the NASA funded and sponsored research

discussed below, an additional topic of interest and relevance is

the agreement which UCSB has negotiated with Digital Equipment

Corp. (DEC). This agreement (see appendix A) provides that UCSB

researchers will conduct joint research in the area of knowledge

based engineering for spatial data base management and use. As

part of this effort, DEC provided one half the costs of a

superminicomputer system for UCSB's use. The VAX 11/750 with a

graphics workstation and peripherals has arrived, been installed

and is operating. The system is dedicated to information

sciences research. We feel that the funding of this effort by a

principal computer science corporate entity is further

verification of the status of UCSB as a center for information

science research.

The material which follows details ongoing work which has

been and is 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

3

with:

* The Conduct of Field Research to Develop a Data Base on The Boreal Forests of North America;

* The conceptual design and recommendations for the phased implementation of a Pilot Land Data System;

* Advanced Data Structures for integration of image and Cartographic Data in a Geographic Information System;

* Applying Artificial Intelligence to Large Networks;

* Review of Literature related to Global Resource Information systems; and

* Knowledge Based Expert Systems for Crop identification.

In the next section, these projects are discussed in greater detail. Ongoing and proposed activities are discussed. This is followed by a brief section summarizing future research directions. This first annual progress report concludes with four appendices containing material relevant to sections of this progress report.

Project Reports

Habitability of the Earth
    D.B. Botkin, K.D. Woods

NASA Pilot Land Data System
    J.E. Estes, J.L. Star, J. Franklin

Advanced Data Structures and Geographic Information Systems
    D. Peuquet

Artificial Intelligence and Large Networks
    T.R. Smith, J.E. Estes, R. Dubayah

Knowledge-Based Expert Systems for Crop Indentification
    T.R. Smith, J.E. Estes, C.T. Sailer, L.R. Tinney

N84
31740

UNCLAS

Habitability of the Earth

Prof. D.B. Botkin and Dr. K.D. Woods

Department of Environmental Studies
University of California, Santa Barbara

Progress Report: Boreal Forest Vegetation Analysis,
Ely, Minnesota

Daniel B. Botkin and Kerry D. Woods
University of California, Santa Barbara

I. INTRODUCTION

The work accomplished during 1983 under the funds from the
NASA Office of University Affairs was part of a larger study
titled "Habitability of the Earth: Analysis of Key Vegetation
Factors." The general purposes of this research are (1) to
develop and test methods to measure land vegetation biomass, net
biological production and leaf area index by remote sensing; and
(2) to apply these measures to estimate the biomass and net
biological productivity of selected biomes, including the boreal
forests and north temperate grasslands. Under funds from the
NASA Office of University Affairs field verification was
conducted in conjunction with remote sensing of pertinent
variables and the development of a data base of relevant material
initiated. Specifically, measurements were made in the Superior
National Forest, Minnesota on those vegetation characteristics
which are being correlated with the following remote sensing
measurements: (a) an 8 band Barnes radiometer mounted in a
helicopter which hovered at 400 feet above each site; (b) the
thematic mapper simulator flown in a NASA C-130 aircraft; (c) MSS
data from Landsat 3; and (d) AVHRR data from the NOAA Satellite.

This report summarizes vegetation data collected as field
verification in the Superior National Forest. Headquarters for
the research were maintained in Ely, Minn., adjacent to the one

1

million acre wilderness of the Boundary Waters Canoe Area. This
location was chosen because (1) within the coterminous United
States, it represents one of the best examples of boreal forests:
(2) the terrain is relatively flat, which is important in initial
tests of techniques; (3) excellent logistics were available,
including jet fuel at the Ely airport, assistance from the US
Forest Service Office, a well maintained road in a corridor
between two sections of the Boundary Waters Canoe Area where
adequate varieties of forest vegetation and stages in forest
development were available; and the existence of good research on
the vegetation of the region.

To review, the primary objective of this research was to
provide detailed ground truth for sites over which intensive
radiometric measurements are taken so as to permit development of
accurate calibration of remote sensors for measurement of
vegetation parameters. There are two steps in the field research
known as (1) plot measurements and (2) dimension analysis. The
goal is to determine the leaf area, biomass and past growth rate
of all trees in a sample plot. In theory, this could be measured
by cutting down all trees in a plot and weighing and measuring
each component. In reality, this is not feasible, expecially
within the budget limitations of our project; but it has never
been feasible in any ecological research project to our
knowledge. As a result the two stages are used to provide the
measures of interest. In dimension analysis, a statistically
valid sample of trees of each species are cut down, cut apart and
weighed by categories (leaves, twigs, branches, etc.) These
weights are then related to dimensions of the trees that can be

2

Habitability of the Earth

obtained rapidly in the field including tree height, tree
diameter, height to the lowest and highest living branches. The
resulting statistical regression equations are then applied to
trees measured for the non-destructive variables on sample plots.
Although these procedures have been done frequently, we have
extended them and improved the statistical validity of the
approach.

The intent, during 1983, was to focus efforts on stands
dominated by two of the major species of North American boreal
forests and the Superior National Forest: black spruce (lowland
stands) and trembling aspen. Single species were chosen because
these species have the broadest geographic ranges of any major
trees in North America, and are major components of the boreal
forest throughout North America. They were also chosen because
in every aspect they represent opposite ends of a vegetation
spectrum: the spruce is evergreen and characteristic of bogs and
other wet sites, and of old growth forests; the aspen is
deciduous and characteristic of dry, upland areas and new forest
stands.

We believe it is especially noteworthy in this report to
point out the funds for the research were not available until
late spring 1983, and that all planning including site selection
for the field research began March 20 with the first trip to Ely
by principal investigators. In spite of this extreme time
constraint, we believe that considerable progress was made toward
the goals of the research.

II. SITE SELECTION AND THE DEVELOPMENT OF A RAPID SELECTION
METHOD

During 1983, basic field procedures involved 1) selection of
sites (plots) for radiometric measurement covering the range of
leaf area index and biomass seen in the region, 2) non-
destructive measurement of a number of simple, biologically
meaningful dimensions on all trees in each sample plot, including
(3) modification of standard forestry techniques to meet the
needs of our research, and (4) destructive sampling of a number
of black spruce and aspen trees for direct measurement of the
parameters of interest (biomass and area by tree component) and
for development of predictive relationships between plot
measurements and these variables.

Progress outlined is of three types: collection of field
data; and refinement of procedures and clarification of
questions.

Even though planning for the project began only three months
before the field season, we have accumulated a very large and
unique data set specifically designed for our project. Plot data
have been acquired for 31 spruce and 31 aspen sites. Plots were
chosen to represent as large a range of biomass and production as
possible.

Initial stratification of these sites by biomass/leaf area
was based on qualitative observations by the principal
investigators. One of the unique contributions of the research
was the development of rapid site selection techniques using a
coordinated team of helicopter and ground crews. In the past,
site selection for ecological research has been done by ground

crews. However, the dense foliage of the boreal forests made
location and selection of sites on the basis of biomass and
productivity extremely slow and inaccurate. The principal
investigators, in conjunction with NASA helicopter pilots devised
the following method: two ground crews of three each with a car
and two way radio were directed to sites .   .incipal
investigators flying in the helicopter. . ..s the   . i ..
could be located rapidly and the fastest and easiest paths to the
sites determined. Moreover, the helicopter crew was able to
locate sites and guide two crews simultaneously. We believe that
this technique has broad potential for application to research
related to the study of the biosphere and Global Habitability in
many areas of the Earth.

III. GROUND MEASUREMENTS MADE ON HELICOPTER REMOTE SENSING SITES

At each site chosen for helicopter remote sensing, a
circular plot of 60 m diameter was mapped, and within this larger
plot five small plots were laid out. On each of these five
plots, standard measurements of vegetation were made, including
the diameter, height, and height to the lowest and highest living
branching. These could then be correlated with the dimension
analysis data described below, and estimates of biomass could be
calculated.

Although initial site selection was strictly qualitative,
based on helicopter and ground observations, the available range
of pertinent variables seems to have been well covered. As a
result, we expect that little additional plot data will be

required in 1984 for the two species, black spruce and aspen, studied in 1983. In addition to the spruce and aspen data, data were also collected for twelve jack pine stands and eight plots with a mixture of two species. Jack pine measurements were taken to determine the accuracy with which two conifers could be differentiated by helicopter remote sensing. The mixed stands were measured because most forests were composed of stands of several to many species, and it was necessary to begin to consider the problem of mixed pixel responses for planning for the 1984 field season. These additional measurements, therefore, should facilitate design of further study involving stands of these types.

IV. DIMENSION ANALYSIS

31 spruce trees and 32 aspen trees, distributed over the size range for these species, were cut down for detailed dimension measurement. Initial analysis of data from aspen trees suggest that, although strong relationships exist between tree dimensions and leaf area, further sampling is necessary to develop satisfactory predictive equations. This is because 1) variance is large, 2) some sampled trees may have been atypical due to unusual weather, and 3) larger samples in some size classes are required. We are also unsure of the biological significance of the relationships seen. Nevertheless, we have been able to use relationships developed from the sacrificed trees to make preliminary estimates of leaf area for aspen plots. These will be compared with radiometric measurements. Biomass estimates for aspen trees and plots are expected be be more

accurate. Similar analyses will commence soon on data from sacrificed spruce trees. We expect that these will proceed more smoothly since both sampling and analysis techniques have been improved through prior experience with aspen.

V. GENERAL COMMENTS

The resulting data set is large -- over 2500 pages of raw data -- and to our knowledge, unparalleled in ecological research. We conducted tests on the accuracy of field measurements through replicate sampling and the consistency was very high. Plot data and data on sacrificed aspen and spruce trees have been checked and verified. We are, as discussed below, attempting to streamline this process for the 1984 summer's work. Foliage and bark samples were also collected and forwarded to LARS and JSC for analysis of reflective and transmittance properties for use in canopy models. Extensive radiometric (TMS) data were collected; all sampled sites were measured at least once and many several times.

We began the field season in 1983 with untested field procedures and _a priori_ assumptions about important variables and relationships, and discovery of previously unsuspected, but important, problems and needs. Plot sampling time has been roughly halved; a four person crew (8-10 workers) was able to destructively sample one to two trees per day, but this rate increased to three trees per day by summer's end. At the same time as rates of data collection increased, accuracy and replicability was improved. We believe that, with no increase in

number of personnel, we can increase the amount of data collected
next year by 50-100% without loss of accuracy or reliability.

At the same time, various problems arose to which we are
currently addressing our efforts. It became apparent early in
the summer that transfer of data from the field to Houston and
its entry into the computer data base and proof-reading was a
major bottleneck. For the coming field season, we are developing
procedures for entering data on magnetic media in Ely, with
immediate proof-reading. Data may then be sent to Houston for
direct entry. By this means the interval between data collection
and beginning of analysis should be decreased by several months.
Initial analyses of 1983 data suggest that forest "background"
(understory plants, shrubs, ground surface, etc.) is important in
determining plot reflectance. We will design next summer's
efforts to better address this problem.

Habitability of the Earth, 1984 Plans

Objectives for 1984 Vegetation Study Field Season,
Ely, Minnesota

2

Habitability of the Earth

Daniel B. Botkin & Kerry D. Woods
University of California, Santa Barbara

The proposed activities for the 1984 field season are
designed to provide additional data needed for field verification
for remote sensing of vegetation biomass, net production and leaf
area. The justifications for this research are given in our
original proposal, and will not be repeated here. Specifically,
the proposed research is to continue to provide the field
measurements to adequately describe the biomass, leaf area, and
productivity for a data base on the boreal forests of North
America. In 1984 we plan to continue work in the Superior
National Forest of northeastern Minnesota and to explore possible
field sites for work in the following years. The list is based
on our understanding of the composition and dynamics of these
forests, experiences of last year, and initial analyses of last
season's data. As analysis proceeds or additional objectives are
considered, this list may change. Field work will be conducted
by a crew of 10, active for 3 months in the Ely area.

1. We expect to supplement last year's data for spruce and
aspen in order to strengthen statistical conclusions and study
year-to-year variability. We will cut-down additional (perhaps
10-15) aspen trees in order to improve statistical predictors of
biomass and leaf area and determine the extent of year-to-year
variability of leaf area in a deciduous species. We are less
likely to need further analysis of lowland black spruce, but we
may sample some trees of upland black spruce to see if
relationships developed for lowland trees apply. We do not
expect to sample additional plots in aspen or black spruce

9

stands, but we may resample a small number of those sampled last year to test replicability.

2. A primary objective is enlargement of our database to include additional species important in the Ely area and boreal forests in general. To this end we sampled 12 jack pine plots in 1983. We will expand the data set for jack pine and collect dimensional and biomass/leaf area data for sacrificed trees. Other important species which may be sampled this summer are paper birch and balsam fir. For each species we would locate 20-30 sample plots and sacrifice a similar number of trees.

3. Since significant areas are occupied by mixtures of species, we must develop techniques for estimation of biomass and leaf area of mixed stands. Last year, we sampled 8 plots of mixed jack pine and trembling aspen. This year we will sample additional plots of this and perhaps other species mixtures. We will use measurements from these plots with radiometric data to study the contribution to the spectral signature of each species in mixed stands.

4. Since an eventual goal is extension of procedures and relationships developed to the entire range of boreal forests, we must test them in other areas. A first step in such testing might be sampling of a small number of sites in another portion of the Superior National Forest (about 100 km from Ely). Relationships between vegetation parameters and radiometric measurements developed for the Ely region could be tested for this region.

5. The next geographical extension of the study (during future field seasons) should be to parts of the North American

boreal forest as remote as possible from Ely. We are looking into possible sites in Maine, Quebec, Alberta, and Alaska. During the summer of 1984 two primary investigators (Botkin and Woods) will make reconaissance trips to two of these areas to make local contacts, inspect facilities, and evaluate vegetation. Such preliminary work will be very valuable for advance planning which should alleviate some of the procedural problems experienced in 1983.

6. Throughout the summer we will be testing and, when appropriate, implementing improvements in sampling and measuring technique, measurement of "background" vegetation, and transfer and verification of data.

Our goals for the end of 1984 are (1) complete field work in the Superior National Forest site; (2) develop statistically valid methods to estimate biomass, biological production and leaf area for spruce and aspen and one or two additional species; (3) obtain an initial understanding of the mixed pixel problem for forest vegetation; (4) determine the accuracy with which several species with similar reflectance characteristics (black spruce and jack pine; aspen and birch;) can be distinguished by remote sensing and biomass, net production, and leaf area measured.

LARGE OR SMALL PART OF BOREAL FOREST MEASUREMENT PROBLEM ?

This is a somewhat heavier work-load than that of last year. However, our 1983 field experience persuades us that, given some prior training of crew (several of the field crew will be veterans of last summer), improved procedures, and better understanding of system and objectives, this load does not represent unrealistic expectations. The work outlined will

provide a database, in conjunction with last year's, of unprecedented completeness for an entire vegetational system. With calibrated remote sensing data the opportunities for detailed and accurate analysis of ecosystem status and function will be highly unique and extremely valuable.

N84
31741

UNCLAS

NASA Pilot Land Data System

Prof. J.E. Estes, Dr. J.L. Star, Ms. J. Franklin

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

NASA Pilot Land Data System

Dr. J.E. Estes, Working Group Chairman
Dr. J.L. Star, Working Group Development Engineer
Ms. Janet Franklin, Graduate Student
University of California, Santa Barbara

Under the sponsorship of NASA Information Sciences and
University Affairs offices, the University Space Research
Association directed a series of workshops intended to develop
the concept and plan for a Pilot Land Data System (PLDS). When
implemented, the PLDS will be the functional equivalent of the
existing NASA Pilot Ocean Data System and Planetary Data Pilot.
These Data base management and delivery systems are in varying
stages of implementation. PLDS is being developed to serve the
same functional needs for the land-related research community.

Members of the PLDS working group include representatives
from NASA Headquarters, the NASA center, universities already
working with NASA, as well as representatives from industry and
the U.S. Geologic Survey. Science and technical subgroups met
several times since summer 1983. UCSB personnel with important
roles in PLDS planning activities are Dr. John E. Estes, Dr.
Jeffrey L. Star, and Ms. Janet Franklin. Dr. Estes is PLDS study
leader and chairman of the science working group. Dr. Star is
working with the technology working group. Ms. Franklin, a UCSB
graduate student, is employed by USRA as project coordinator for
this effort. She has spent the time since September 1983 in
Washington at the USRA offices. Other UCSB staff and students
have contributed to the planning effort as well. The final
meeting of the PLDS Working group was held in the Winter of 1984
at NASA Ames Research Center to bring this phase of the work to a

conclusion and begin editing a final report. Dr. Caldwell McCoy
and Shelby Tilford along with other NASA Headquarters personnel
were briefed on the results of the Working Group's deliberations
in early Spring 1984. The report of the Working Group was
published as NASA Technical Memorandum 86250: "The Pilot Land
Data System: Report of the Program Planning Workshops in July
1984. A copy of the report is attached.

The purpose of the PLDS is to improve the ability of NASA
and NASA-sponsored researchers to conduct land research, by
developing a minimum but representative set of data archive,
management, and processing capabilities. These capabilities will
be designed to satisy a majority of the information systems needs
of NASA-sponsored land research community (including researchers
inside and outside NASA), and should serve as a proff-of-concept
demonstration to provide a basis for future system expansion and
development.

Principal guidelines for PLDS development include:

*   Coordination with appropriate science users during design
and implementation is vital.

*   Discipline scientists must specify their requirements,
and the information scientists must then base their designs on
these needs.

*   System design and implementation must be phased and
modular.

*   Data bases are most viable when maintained by researchers
with a long-term commitment to use of the data.

The FY84 goal is to establish the concept and a preliminary
program plan. NASA Tech Memorandum 86250 and subsequent meetings
between NASA Center and Headquarters personnel are continuing to
produce excellent progress towards this goal. Principal

development activities are expected to begin in FY85, with pilot
demonstration completed in three to five years. Development is
to be in response to science requirements.

The scientists involved in the PLDS development have taken
primary responsibility for specifying requirements. This task
began with development of a series of operational science
scenarios. The science scenarios developed to date include:

* Vegetation Biomass and Productivity
    Botkin (UCSB)

* Terrestrial Ecosystems and Productivity
    Brass et. al. (NASA/Ames)

* Land Surface Climatology
    Price, Middleton (NASA/Goddard)

* Large Area Landcover Mapping
    Estes, Franklin (UCSB)

* Soil Erosion and Hydrological Modeling
    Regen (UM), Langran (NASA/NSTL)

* Sedimentary Basins
    Farr (JPL)

* Regional Mineral Resource Assessment
    Albert (USGS)

* Monitoring Environmental Change
    Albert (USGS)

These scenarios have been analyzed to determine data sources,
processing capabilities and throughput needs, and information
output products, to evaluate the overall requirements of the
PLDS. The scenario approach has been taken to be able to focus
on specific tasks the scientists would like to perform, and then
step back and consider the generic functions and structure of a
pilot implementation. By selecting scenarios from a range of
topic areas, we are able to identify the basic similarities of
data and analysis requirements, which are then used to drive the

specifications of a minimum operational system.

At the present time, it is envisioned that the science subgroup should continue in both an operational mode, by exercising the system in the conduct of research, and in a review and advisory role. These activities must continue during pilot operation.

The system developers will design and implement PLDS hardware and software in response to the scientists' requirements. The system developers will design and implement PLDS hardware and software in response to the scientists' requirements. The system developers will also provide operational support during the pilot demonstration period.

Configuration and technical onsiderations include:

* Pilot Operation and Administration

* Data Management
    Data Dictionaries, Data Base Management, Data Access and Storage, Browse Facilities

* Processing
    Software Libraries, Distributed Processing, Unique Hardware Facilities.

* Communication
    Network Control and Accounting, Traffic Flow, Standards, and Possible Configurations

* User Interface
    User Requirements and Capabilities, Accounting, Local Area Networks, Standards/Guidelines, Workstations.

NASA Center Working group Members have met along with Headquarters Personnel have met. Scenarios dealing with Land Surface Climatology and Sedimentary basins are being considered for initial Pilot development testing. Center roles in Technology development are being refined, and levels of funding

are being discussed, along with the overall structure for the
pilot. UCSB personnel hope to continue to assist and participate
in the development, demonstration and implementation of PLDS by
the Information Systems Office at NASA Headquarters.

N84
31742

UNCLAS

Advanced Data Structures and Geographic Information Systems

Dr. D. Peuquet

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

Advanced Data Structure and Geographic Information Systems

(805) 961-3772

Dr. Donna Peuquet
University of California, Santa Barbara

Paul Smith —
—> Bill Campbell
9541.

Introduction -

Large integrated software systems known as Geobased or
Geographic Information Systems (GIS) have been developed in the
past fifteen years which can perform all phases of storage,
maintenance, analysis and retrieval for geographic data. GIS's
based on shared, large-scale, integrated data bases are used for
a wide range of applications which deal with geographic data for
government agencies and private corporations worldwide. They
have already become an indispensable tool for managing
biophysical, biological and socio-economic data in a timely and
efficient manner. However, only a limited volume and range of
data types can be handled by any one system. ... This using
present GIS technology. Overcoming these two prob... ...
increasing data volume and range of data types,critical for the
analysis of complex environmental problems if we are to
successfully manage our resources on local, regional, and global
scales [Estes, 1980].

The purpose of this project is to begin to examine the
current state of the art in specified areas of GIS technology and
to establish a common ground for research. Study of the question
of very large, efficient, heterogeneous spatial databases is
required as NASA begins to explore the potential application of
remotely sensed data for studying the long-term habitability of
the earth. In addition, neither GIS nor remote sensing
technology can achieve their full potential until they are

1

Advanced Data Structure and Geographic Information Systems

Integrated. The research proposed herein is a significant step in this direction.

Recent research by Dr. Donna Peuquet (Department of Geography, UCSB) has examined the tradeoffs between raster and vector processing of geographic data and concluded that neither can ever be adequate by itself for large-scale integrated databases. Further, the need for raster-to-vector and vector-to-raster conversion processes, which represent a significant system overhead, must be avoided whenever possible. Initial work on a hybrid spatial data structure which incorporates characteristics of both raster and vector data structures was subsequently performed [Peuquet, 1983]. This structure, called the VASTER structure, represents a first step in the development of new approaches to geographic data handling for GIS to handle integrated databases of many types of geographic data on a global scale.

A similar point of view on the part of researchers at The National Aeronautics and Space Administrations (NASA) Goddard Space Flight Center (GSFC) has resulted in the independent development of another structure termed the Topological Grid Structure. It is important at this point in time to coordinate the development of these two new types of data structures for two reasons; 1) to determine the relative performance characteristics of each structure for various applications and operationa environments so that their relative merits are known for future potentia real-world GIS applications; and, 2) to provide direction for further research in this area. As such UCSB personnel under the direction of Dr. Donna Peuquet is

2

Advanced Data Structure and Geographic Information Systems

engaged in a research project with GSFC personnel which is being
conducted in several phases each containing a number of research
tasks.

Phase I of this project (9/1/83 - 1/31/84, $14,913.00)
involved three tasks:

1) a review of spatial data structures and general
approaches to storing spatial data;

2) preliminary work on the development of a set of
primitive operations which are required of any geographic
information system (GIS); and,

3) preparation specifications for a testbed system which
will subsequently be used to compare the Vaster data structure
developed by the principal investigator with the Topological
Raster Structure developed at NASA/GSFC.

The Phase I work was explicitly intended to explore the
theoretical context of comparative data structure and algorithm
evaluation, and to derive detailed compatibility requirements, as
a preliminary step before actual testing can take place.

Phase II (2/1/84 - 12/31/84, $38,581.00) constitutes the
most critical phase of this work. The primary emphasis of the
proposed second phase of this work will be to construct the
Vaster testbed system which will be specifically tailored
for comparability with the Topological Raster Structure testbed
system. This work will consist of major extensions to tasks 2
and 3 above:

Task 2

The preliminary set of GIS primitive operations developed in
phase one will be refined and finalized for the purpose of this
research. This includes categorizing procedures into
functional groups.

3

Advanced Data Structure and Geographic Information Systems

ORIGINAL PAGE IS
OF POOR QUALITY

Task 3

The development of the testbed system for the Vaster

Structure will be completed. Specific elements being addressed

in phase two are:

a) Mutually agreed selection of "real world" data sets
which are to be used for the comparative tests. These data sets
are to be supplied by NASA/GSFC,

b) Implement.   or test
procedures which will be the basis of the comparative tests. The
                                          . . ....ur are being
. . ., ..reement between the NASA/GSFC contract
representative and the principal investigator. Algorithms for
each procedure must be carefully selected for implementation in
the two testbed systems to insure their comparability; and,

c) Initial test runs of the Vaster testbed system will be
made in order to determine the amount of machine resources needed
to perform the actual comparative analysis computer runs.

This work utilizes a VAX 11/750 system currently being

acquired through an equipment grant from DEC. This system, which

was installed this past spring, is dedicated exclusively to

information sciences research particularily in the areas of

spatial database and knowledge-based geographic information systems.

As can be seen above this work funded by NASA/GSFC will

continue into the coming year. This grant will continue to serve

a role integrating this work with other research being conducted

at UCSB and providing support to this work in terms of personnel

and equipment as needed. We feel the synergism created by the

integration of this effort within the overall framework of this

Office of University Affairs projects significantly benefits both

Advanced Data Structure and Geographic Information Systems

projects.

N84
31743

UNCLAS

Artificial Intelligence and Large Networks

Prof. T.R. Smith, Prof. J.E. Estes, Mr. R. Dubayah

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

"APPLYING ARTIFICIAL INTELLIGENCE TO LARGE NETWORKS"

Dr. Terence R. Smith, Dr. Jeffrey L. Star, Mr. Ralph Dubayah
Department of Geography
University of California, Santa Barbara

The Problem

Personnel of the Information Sciences Working Group at UCSB
are actively involved with researchers at the NASA Ames Research
Center, investigating the use of artificial intelligence in
networked information systems. The problem deals with networking
NASA centers and associated researchers in more efficient ways.

The ultimate goal is to be able to access and process
geographically dispersed data sets with the use of geographically
distributed software. This requires capabilities in the areas
of: distributed data base management, long-haul networking,
information presentation, distributed problem solving and
artificial intelligence, to be combined in new and innovative
ways.

At this time, a number of major NASA initiatives revolve
around the problems of data accessibility and transfer of
information. In particular, the Pilot Land Data System, the
Pilot Ocean Data System, and the Pilot Climate Data Base
Management System (as well as a proposed Global Resource
Information System) are structured as large computer networks.
The goals of these pilot efforts include reducing the difficulty
of finding and using data, reducing data processing costs, and
minimizing incompatibility between different data sources for
researchers both within and outside NASA. The techniques of
artificial intelligence have application to this problem. In

1

particular, AI search procedures, knowledge representation and
control strategies are relevant for this task. Researchers at
UCSB are focusing on how these AI techniques may be employed in
distributed problem-solving, in which both data and procedures
are distributed throughout some network.

The Research

A future system, in order to respond to a user query might
require computations on several databases at one location using
software and hardware at a second location. The major
aspect of such a system is an AI control system with knowledge
about the databases, software and hardware throughout the
network, and able to co-ordinate the (possibly parallel)
processes required to satisfy the user's query.

In particular such a system might include the following
knowledge bases:

1- Meta - Knowledge Base - the system's knowledge about itself.

2- Network Knowledge Base - knowledge about the network and
transfer of information through the network.

3- Database Knowledge Base - knowledge about the various
distributed databases.

4- Image Analysis Knowledge Base - knowledge about the analytic
and processing capabilities of the various network nodes.

5- Production Knowledge Base - knowledge about problem solving to
satisfy a user's specific query.

6- Learning Knowledge Base - knowledge about how to improve the
upgrade of the system.

Each of the knowledge bases is a topic of research in

Artificial Intelligence and Large Networks

Itself, as is the overall control structure. For a particular
application, expert domain knowledge bases can be added (i.e.,
mineral discovery, geotectonics, vegetation discrimination,
biomass mapping, change detection). Finally, the research will
incorporate results of state-of-the-art AI research concerning
distributed problem-solving.

Relation to Other Research

One basis for this work is research currently being carried
out at UCSB under the guidance of Terence R. Smith, Donna J.
Peuquet and John E. Estes which relates to the application of AI
in building a knowledge based geographic information system
(GIS). This work is currently funded by U.S. Geological Survey,
National Science Foundation, and Digital Equipment Corporation. As
stated in the introduction these researchers have just acquired a
VAX 11/750 for dedicated AI research, in cooperation with Digital
Equipment Corporation researchers.

Future Research

The research with NASA-Ames has important implications for
future image and data processing. It is directly addressing one
of the most difficult data-processing problems faced by NASA. To
effectively and efficiently link the diverse spatial data
processing capabilities of a geographically dispersed user
community. A full solution to the problem will undoubtedly
require much further research that will both draw upon and
contribute to our knowledge of now AI may be used to control
difficult data processing problems.

3

N84
31744

UNCLAS

Knowledge-Based Expert Systems for Crop Identification

Prof. T.R. Smith, Prof. J.E. Estes, Ms. C.T. Sailer

Remote Sensing Reserach Unit, Department of Geography
University of California, Santa Barbara

Mr. L.R. Tinney

E.G. & G., Las Vegas, Nevada

Knowledge-Based Expert Systems for Crop Identification.
Terence R. Smith, John E. Estes, Charlene T. Sailer
University of California, Santa Barbara
and
Larry R. Tinney
E.G. & G. Las Vegas, Nevada.

For the past year, UCSB has been conducting a cooperative study of the application of expert systems in the interpretation of digital remotely sensed imagery. Funding for this project was provided by the National Aeronautics and Space Administration Office of University Affairs and the California Space Institute for the 1982-1983 fiscal year.

California Space Institute (Calspace) is a multicampus research unit of the University of California. Calspace funding represents "seed " money to support basic research in the space applications area. Funding for this project went to support graduate student research in the applications of expert systems in image analysis, including the research of doctoral student Charlene Sailer. Ms. Sailer is presently a NASA Fellow continuing work in this research area in cooperation with the Johnson Space Flight Center.

Research being conducted in this area is exploring the methodology and components of human image analysis in an agricultural applications area. The long term goal of this research effort is the development of an improved understanding of the interactive man-machine environment. In such an environment, as many feature inputs as practical would be automatically derived from a data-base and input into an expert system decision-making procedure. This procedure could then

provide "expert" assistance to a trained image analyst,to upgrade
and improve the quantity and accuracy of the information
extracted from the input data. The research being conducted for
this project are the first steps towards this long-range goal.

To date, in this project, all aspects of the image
interpretation process are being explored, beginning with the
fundamental elements of the human image interpretation process.
This is being done with the goal of achieving an improved
understanding of these elements and their significance in the
decision-making process. A comparision of the similarities and
differences between manual and automated image interpretation
techniques as currently practiced is also being explored. This
effort is being accomplished to better understand current areas
of overlap and to better define those intersections of human
cognitive skill and machine technological capabilities wherein
manual and automated procedures may eventually meet.

Several papers have already been generated from this work.
These papers seen in Appendix C were presented at the 17th
International Remote Sensing Symposium at Ann Arbor, Michigan at
the National Telecommunications Conference in San Francisco,
California. In addition to these papers a number of preliminary
experiments have been conducted. These experiments involved
sessions where image analysts were shown two sets of multi-date
LANDSAT imagery, and a set of LANDSAT images of predominately
agricultural areas which exhibited varied agricultural practices.
The analysts then interpreted these image sets with an
interviewer who queried the analysts on points of the
interpretation.

Expert Systems

Future work in this research area will be conducted by Ms. Sailer for her doctoral dissertation through the partial support of NASA - Johnson Space Flight Center. Her dissertation will explore the growth cycle of crops in the San Joaquin area of California employing multi-spectral, multi-date LANDSAT imagery. Ms. Sailer will be working with expert image analysts at JSC to develop algorithms simulating the decision logic used in the discrimination of crops. The final product for this research will be a series of AI linked computer programs which will assist an image analyst by providing expert advice and the probability of an area being an agricultural area and the probability of occurance of a certain crop in that region.

Summary

Prof. John E. Estes and Dr. Jeffery L. Star

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

Researchers at UCSB are in the initial phases of building an integrated research program in information sciences around a number of key themes of interest to NASA. These themes are an outgrowth of a thorough analysis of the current state-of-the-art in the areas of: remote sensing algorithm development, spatial data bases, artificial intelligence, and global scale studies. As seen in the preceding section projects currently underway involving these themes include: Boreal Forest Vegetation Analysis, NASA Pilot Land Data System, Advanced Data Structures for Integration of Image and Cartographic Data in Geobased Information Systems, Applying Artificial Intelligence to Large Networks, Global Resources Information System, and Knowledge Based Expert Systems for Crop Identification.

Table I shows the intersection of these projects with the major research themes. Information Science Working Group personnel are actively pursuing research in each of the primary areas of research interest listed in our original proposal. We feel that in each of these major areas we are making progress towards a deeper understanding of a variety of significant research issues.

During the discussions with NASA Office of University Affairs personnel which led to the funding of the Information Science Working Group at UCSB, two general objectives over and above actual research objectives were brought foreward. It was agreed that this Grant would be used as a focusing mechanism to help support the integration of on-going research efforts at UCSB. In addition, the Grant would also be used as a catalyst to

encourage and attract cooperative research between other NASA
Divisions and NASA Centers and if possible between NASA and other
Federal Agencies. Again, based on our work to date, we feel we
have made significant progress towards meeting these objectives.
Indeed, research discussed in this report involves the Earth
Science and Applications and Life Sciences Divisions, the
Information Sciences Office, and of course, the Office of
University Affairs.

This work also entails major cooperative research efforts
with personnel from NASA's Ames Research Center, Goddard Space
Flight Center, Jet Propulsion Laboratory, and Johnson Space
Center. In addition personnel from the National Science and
Technology Laboratory (NSTL) at Bay St. Louis, Mississippi are
involved with UCSB researchers in the Pilot Land Data System
effort discussed above.

During the period May 1, 1984 to April 30, 1985, UCSB
researchers will continue the efforts as described above. Funding
in the coming year provides for: continuity of funding of key
staff, continuation of research activities with NASA Headquarters
and Center Personnel, field research in the Boreal Forest, and
matching funds for the Digital Equipment Corporation (OEC) VAX
11/750 which was provided to UCSB.

With respect to the last item, UCSB is very pleased that the
Office of Information Science and the Office of University
Affairs have agreed to provide a small portion of the matching
funds for this equipment. We feel that the multiplier effects
for this project will be extremely significant. This

superminicomputer system is solely dedicated to information

sciences research.

Finally, we at UCSB appreciate the opportunity provided by

this Grant from the NASA Office of University Affairs. We feel

that the area of information science represents a real challenge

as we move towards the application of satellite remote sensing to

global studies. The ability to use information in an effective

and cost-efficient manner, derived from the globally consistent

data acquired by NASA and other satellites, to combine these data

with aircraft and a variety of other scientific data/information

sources, is vitally important. Important as we attempt to

understand questions involving the interaction of climate,

atmosphere, hydrosphere, lithosphere, and biosphere, research

conducted under the auspices of this Grant is directed towards

this end.

Budget

PROPOSED BUDGET

AGENCY: NASA   AGENCY #: NAGW-455
P.I.:   J. ESTES
PERIOD: 5/1/84 to 4/30/85
TITLE: NASA Remote Sensing Information Sciences Research Group

SALARIES*

ESTES, J., Principal Investigator
    1 smr mo @ 100% @ 5433   5,433

SMITH, T., Professor
    1 smr mo @ 100% @ 4504   4,504

PEUQUET, D., Asst. Researcher
    1 smr mo @ 100% @ 4278   4,278

STAR, J., Assoc. Dev. Eng.
    2 mos @ 25% @ 2521   1,261
    10 mos @ 25% @ 2773   6,933

TBD, Staff Research Assoc. I
    2 mos @ 50% @ 1484   1,484
    3 smr mos @ 100% @ 1632   4,896
    7 mos @ 50% @ 1632   5,712

TBD, Research Assistant
    5 mos @ 24% @ 1494   1,793

TBD, Senior Typist Clerk
    2 mos @ 25% @ 1126   563
    10 mos @ 25% @ 1239   3,098

SALARIES TOTAL: 39,955

BENEFITS:

Estes, J.
    5433 @ .2262 (w/SS)   1,229

Smith, T.
    4504 @ .2262 (w/SS)   1,019

Peuquet, D.
    4278 @ .33 (actual)   1,412

Star, J.
    5421 @ .31 (1984)   1,681
    2773 @ .32 (1985)   888

Staff Research Associate
    8828 @ .31 (1984)   2,737
    3264 @ .32 (1985)   1,045

BENEFITS (Cont.)

Research Assistant
    1793 @ .0050 (yr)   9

Senior Typist Clerk
    2422 @ .31 (1984)   751
    1239 @ .32 (1985)   397

BENEFITS TOTAL: 11,168

EQUIPMENT

Partial funding for purchase of VAX 11/750   15,000

Interactive workstation   3,000

High-speed modem for workstation   1,000

EQUIPMENT TOTAL: 19,000

TRAVEL

1 trip for 2 people for 3 days from
    Santa Barbara to NASA Headquarters
    1 RT airfare for 2 persons   1,200
    6 days of per diem @ $75/day   450
    Car rental and transportation   100
                                   1,750

1 trip for 2 people for 3 days from
    Santa Barbara to NASA/JSC
    1 RT airfare for 2 persons   428
    6 days of per diem @ $62/day   372
    Car rental and transportation   100
                                   900

1 trip for 3 people for 2 days from
    Santa Barbara to NASA/ARC
    1 RT airfare for 3 persons   444
    6 days of per diem @ $62/day   372
    Car rental and transportation   100
                                   916

TRAVEL TOTAL: 3,566

SUPPLIES

Office supplies	1,000
Office equipment rental-word processor	1,000
Telephone tolls & equipment	1,750
Xerox & Library copying	500
Photographic supplies	750
Computer charges (Itel AS/6 & VAX 11/780)	3,500
Report Preparation	600
SUPPLIES TOTAL:	9,100

TOTAL OF DIRECT COSTS (Estes):	82,789
TOTAL OF DIRECT COSTS (Botkin):	48,294
DIRECT COSTS TOTAL:	131,083
**OVERHEAD OFF-CAMPUS RATE 25.8%: (on $112,083)	28,917
TOTAL:	160,000

* Salaries calculated: Current rate X 11% (academic) or 10% (staff) range adjustment
  effective 7/1/84. Merit increases are added according to the appropriate title
  schedule.

**Indirect cost rate of 25.8% (off-campus) is effective through 6/30/84,
  provisional thereafter.

ADDENDUM TO PROPOSED BUDGET

AGENCY: NASA
P.I.: D. Botkin
PERIOD: 5/1/84 to 4/30/85
TITLE: NASA Remote Sensing Information Sciences Research Group

BWCA, Ely MN, Field Work Budget

SALARIES

D.B. Botkin, Professor
    1 smr mo @ 100% @ 5459/mo   5,459

K.D. Woods, Assistant Spec. II
    3 smr mos @ 100% @ 2037/mo   6,111

TBD, Lab Helpers (5 persons)
    1 mo @ $5.51/hr @ 833 hrs   4,590
    3 smr mos @ $6.06/hr @ 1667 hrs   10,102

TBD, Assistant I (5 persons)
    1 mo @ $5.50/hr @ 833 hrs   4,582
    3 smr mos @ $5.97/hr @ 1667 hrs   9,952

SALARIES TOTAL: 40,796

BENEFITS

Botkin
  5459 @ .1612%   880

Woods
  6111 @ .31%   1,895

Lab Helpers
  14,692 @ .31%   4,555

Assistant I
  14,534 @ .0115%   168

BENEFITS TOTAL: 7,498

        DIRECT COST TOTAL: 48,294

      INDIRECT COSTS @ 25.8%:   12,460

               TOTAL: 60,754

Appendices

A. Knowledge Based Engineering for Spatial Database
Management and Use
     D. Peuquet

B. Annotated Bibliography of Global Resources Information
     J.E. Estes, J.L. Star, M.J. Cosentino, L.J. Mann

C. Applications of Artificial Intelligence to Remote
Sensing
     L.R. Tinney, C.T. Sailer, J.E. Estes

N84
31745

UNCLAS

LN84 31745

Appendix A.
Knowledge Based Engineering for Spatial Database Management and Use

Dr. D. Peuquet

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

EXTERNAL RESEARCH PROGRAM SUMMARY SHEET

TITLE OF PROPOSAL:

Knowledge Based Engineering for Spatial Database Management
and Use

PERIOD OF PROPOSAL:

1 February 1984 to 31 March 1985

NAME OF UNIVERSITY:

University of California, Santa Barbara

UNIVERSITY CONTACT:

Marguerite McIntire
Research Development and Administration
University of California at Santa Barbara
Santa Barbara, CA 93106
(805) 961-4036

PROJECT CONTACT:

Dr. Donna Peuquet
Department of Geography
University of California at Santa Barbara
Santa Barbara, CA 93106
(805) 961-3772

DEC COORDINATOR:

Mary Kay Lawrence

DEC PROJECT SPONSOR:

Bruce Palmer

Appendix A

Motivation for DEC Sponsorship -

CAEM is interested ink this project as we are
undertaking a project in the GIS and knowledge-based GIS areas as
a critical part of our strategy in the Earth Resources
Engineering market space. UCSB is concentrating on advanced
development that we would otherwise not be able to address in the
early stage of our project due to resource constraints. Of most
interest to us is the query language and learning mechanisms they
are doing, both specific to GIS.

Background -

Large integrated software packages known as Geographic
Information Systems (GIS) have been developed in the past fifteen
years which can perform all phases of storage, retrieval,
manipulation and reorganization of spatial data. GIS based on
shared databases are used for a wide range of applications
involving fundamental research, as well as government activities
and private corporate behavior. These applications range from
urban planning on a local scale to global resource management.
GIS have already become an indispensable tool for managing
geologic, biologic and socioeconomic data in a timely and
efficient manner, and for facilitating the use of these data for
analysis and decision-making.

Many problems of geographic analysis require the
storage, retrieval, manipulation and reorganization of databases
that:

   a) contain many diverse data types, such as real-valued
      functions (e.g., digitized elevation data, rainfall data),
      vector-valued functions (e.g., land use data) and functions
      taking general symbolic values (e.g., place-name data).

   b) contain many layers of data of different types,
      interrelated by complex (implicit) relationships.

   c) contain very large numbers of data points.

   d) are employed in solving a large array of different problems.

The system will be developed to express both objects
and spatial relations in terms of a logically-based spatial
language (a subset of predicate calculus). Since many of the
"reasoning" and "learning" procedures in the system depend on the
application of heuristic knowledge, we term our system a
knowledge-based geographic information system (KBGIS).

2

Research to date has indicated the promise of our
approach, and has given rise to a large number of important
problems requiring much more detailed investigation. The support
that we are requesting will be used to pursue several of these
problems both theoretically and practically, and to incorporate
the results into our prototype system. We believe that our
research will result in dramatic improvements in the capability,
efficiency and versatility of GIS for all potential application
areas.

An Overview of the Proposed Research -

Preliminary research by the investigators into the
construction of a knowledge-based GIS has indicated that recent
developments in AI, database and spatial data processing
techniques, although developed by independently evolving
technologies, can be combined and applied in the construction of
a new and quite distinct type of GIS which will overcome, to a
significant degree, many of the inadequacies exhibited in current
GIS. For example, much research in AI is based on the
development of knowledge-based (or heuristic) search procedures.
These can be combined with tree-based storage and retrieval
techniques developed within the field of database management and
the recent developments in tree structures and algorithms for
spatial data applications. This combination could potentially
yield a system with data volume capacities, levels of efficiency,
and ranges of applications well beyond those that each body of
knowledge alone could provide.

The integration of these multiple technologies into a
single system to handle geographic data must be done
simultaneously so that the various capabilities provided by each
can complement those of the others and give rise to a single,
cohesive system. Since the construction of such a system has not
yet been attempted, the methods of integration will themselves be
a topic of the proposed research.

Preliminary research, funded by the U.S. Geological Survey, is designed to demonstrate that AI techniques can be used in a GIS to deal with practical geographic problems. The work consists primarily of building a small "proof-of-concept" system. This system is currently restricted to answering queries in a given domain. This domain centers around the use of remotely sensed data for natural resource problems. In the process of carrying out the initial work, a number of areas requiring basic research have arisen which are beyond the scope and duration of the initial proof-of-concept project. We believe that the proposed current research is even more important than the proof-of-concept research in terms of its potential impact on GIS technology.

The proposed research is intended to investigate several fundamental questions which have arisen during the course of our preliminary investigation. These new questions encompass four areas:

1. Questions concerning the performance of and modification to the database structure

2. Questions concerning the definition of spectra in quadtree structures and their use in deriving search heuristics

3. Questions concerning extension of the knowledge-base representation and the associated search procedures

4. Questions concerning the concept learning algorithms

Deliverables/Milestones -

1. First demonstrable system which can answer a limited range of queries and update limited components of the database December, 1983.

2. Rule-learning component and autonomous database modification functional June, 1984

3. Concept learning and query language functional July, 1984.

4. Spatial spectra and search heuristics functional October, 1984

5. Complete Demonstration system functional December, 1984

6. Coding for prototype system begins January, 1985

7. Submission of technical report on demonstration system March, 1985

The completion of the demonstration system is also considered to represent the completion of the currently proposed

phase of the project. Upon completion of the demonstration
system, a hands-on demonstration of its capabilities will be
provided for DEC technical personnel at UCSB or at the DEC
sponsor's facilities in Marlboro, Mass., whichever is mutually
determined to be most appropriate.

The technical report on the demonstration system will
consist of a technical description of the primary components of
the system and an in-depth discussion of the knowledge gained via
its construction. Discussions of the general concepts associated
with specific aspects of the system will also be prepared for
publication in scholarly journals during the course of this work.

Risk Assessment -

There is no large r'   .ct.
The proposed research represents the first attempt to integrate
multiple new technologies : .
of this, no marketable software products ...
needed from this project. However, knowledge will be gained
concerning the basic principles involved and this knowledge will
be of specific use in a current DEC project. It should be noted
that the working LISP is essential for this project. UCSB has
agreed to be a DEC LISP field test site and any problems with
that product will cause a slip in their schedule.

Value of Proposed Research -

DEC Techical Evaluation group estimates that a minimum
of two man-years could be saved in development of similar
functionality. This research will have direct bearing on the
CAEM's Earth Resources market strategy for products of this type.
We feel the research warranta a Vax 11/750 class system. The
ability of DEC to use the results of this research will be
greatly enhanced if the work were done under VMS and with a DEC
version of LISP. UCSB originally intended to use UNIX and Franz
LISP to do this project. Therefore, CAEM is requesting that in
addition to the 50% grant against the purchase of the 11/750
configuration, that DEC grant a 100% allowance for the memory and
software to run VMS and LISP. Additional memory is needed due to
the unusual memory requirements for DEC LISP.

Current Status -

A research team headed by the investigators, is
currently partially funded by the U.S. Geological Survey to build
a simple proof-of-concept system. This system will have a
limited range of queries and will also be restricted in the size
and variability of the database.

Support for the expanded work has also been requested
from the National Science Foundation. This additional support is
to be used primarily to provide salaries to the investigator for
his and graduate research assistants for time devoted to this
project.

Technical Status -

DEC and UCSB agree to meet on a quarterly basis to
obtain project status, demonstrations of functional software and
discuss technical problems and solutions of the project. UCSB
providing an account and modem on the research machine will
enable Digital employees to send and receive electronic mail from
UCSB personnel.

N84
31746

UNCLAS

Appendix B.
Annotated Bibliography of Global Resources Information

Prof. J.E. Estes, Dr. J.L. Star,
Mr. M.J. Cosentino, Ms. L.J. Mann

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

Appendix B

An Annotated Bibliography of Global Resources Information

Systems Related Publications

by.

John E. Estes. Jeffrey L. Star, Michael J. Cosentino. Lisa J. Mann
University of California. Santa Barbara

OSTA Data Systems Planning Workshop - DOCUMENTS

User Requirements in Support of OSTA Data Systems Planning
Workshop. Workshop Document #1

This document is a condensation of the 13 ecosystems
discipline requirements. Included for each of the 13 disciplines
are sections entitled: workshop package description. discipline
description with its hierarchy and timeline detailed. and summary
of requirements.   Additional sections include a discussion, on
the use of charts. that display the relationship of platforms and
sensors to natural phenomena measured and to specific data bases
which contain the data.

Anthony Villasenor Data Systems Survey
Systems Division, Goddard Space Flight Center
Workshop Document #2 August 22. 1979

Data system capabilities and overview of specific goals
for each discipline.   A detailed discussion is included for each
discipline's research experiment and contains the sensor to be
used. products and processing fees for each discipline.
distribution of the data. on-board procedures. space/ground
requirements and assessments.

Technical Support Package to Data Systems and Disciplines
Panels. Attachments to OSTA Alternative Data Systems Concept
Workshop Document #3 September 11. 1979 5 Attachments

1) Dr. J.H. McElory. Communications Division NASA Headquarters
   Presentation to the Space Sciences Board's Committee on Data
   Management and Computation on the OSTA Data System.

Today's environmental and resource problems are so
complex that no one discipline can solve them.   The OSTA data
system has to be transparent in meeting the users' needs and

include data collection to develop models and understanding of
multi-disciplines of the atmosphere and the earth. Solving the
problems include a discussion of processing vs. cost/time. data
base management, data delivery, and user access.

2) D.A. Krueger, Chief Systems Division
   A Data System Concept for Application Missions
   August 1979

   A system concept that includes an architecture that
produces simple. easy to implement. technical interfaces: that
does not require major organizational changes to achieve an
approach that allows one user to change data requirements with
out attracting others: economic benefit: an architecture that is
modular: and a system model that can be used to organize what is
done.

3) Fred C. Billingsely, Jr.
   Data Base Systems for Remote Systems. June 1979

   Landsat and other satellite are returning ever-
increasing amounts of data which must be validated. calibrated.
and accurately geographically located to be of highest value.
Formats and cataloging philosophies must be formulated to
maximize compatability of data archives since data are being used
operationally and the archives are rapidly growing. Decisions
are needed to assure that the growth of data archives are orderly
and that dta generated from various sources can be interrelated
and used by a wide variety of analysts.

4) Data Base Management Systems Panel Workshop: Executive
   Summary JPL, August 1979

   The concerns of DBMS operations include cataloging.
acquiring. managing. storing, archiving, accessing, and
disseminating all data for an application. A DBMS must also be
able to provide various levels of support to research activities
and provide both manipulation and analytic functions. From the
panel discussion and their status and future needs for earth
observations data management. recommendations were made. These
include a central browse facility. data archiving. spatial
information system. data base and data format standards for data
sharing. improvement of data networking capabilities. identify
needs for non-NASA data, and identify a comprehensive data
management program.

5) Geophysical Data Interchange Assessment 1978
   Report by the Committee on Data Interchange and Data Centers
   of the Geophysics Research Board

   The growing concern about the environment is producing
a corresponding increase in the flow of geophysical data.
along with recommendations to accommodate the growth of data. In
planning data-intensive projects the funds necessary for data
archiving and for long-term preservation and distribution must be

included from the onset: geophysical observations made regularly
by Federal agencies and scientists are not deposited in WDC's but
in data centers provided to house the necessary resources and to
take custody of the data to avoid loss or damage: a series of
contacts with WDC's in other countries used to reach agreement on
common formats for related data, to ensure more timely data
exchange, to expand the types of data exchange, and to explore
digital data links: and duplicate copies of all physical data is
recommended to protect against possible damage. Appendices
include reports from atmospheric sciences, glaciology,
oceanography, solar-terrestrial physics, solid-earth geophysics,
and space science data.
    Supplement to the OSTA Alternative Data Systems Concepts.
    OSTA Concept Study Workshop Document #4. September 1.. 1979

        Concept study includes block diagrams, evaluation
criteria, and data system options evaluation. 37 data system
implementation options are identified on the data system block
diagrams. Each of these are evaluated and documented to include
a description, justification, quantitative and qualitative
evaluation, address problems, cost, user accessibility, and
provide a summary.

        Technical Support Package for Data System by Discipline
        Panels - Workshop Document #5
        Supplements to the Data System Survey #5.
        Prepared for the OSTA Data Systems Planning Workshop
        September 11, 1979

        Supplement1. ADS suggested to develop an overall
modular architecture for organizing OSTA data sets and data
systems, to develop ADS in a NASA-user oriented access, developed
in an evolutionary way, and to support data access in 1985-1995
time frame to reduce barriers to computer based information
exchange between OSTA and other users. Supplement 2. Fast
facts of the DIDS. This system presents users with a varied menu
of geographical and data categories with special features such
as, a zoom option to isolate specific areas on a map, and data
display in shades of color on maps of the area of interest.
Supplement 3. Global Weather Experiment, a report of the FGGE
Advisory Panel. Goal of the panel is twofold, to acquire a
detailed body of observational data on the behavior of the global
atmosphere and ocean, and ultimately to acquire better knowledge
and understanding of the atmosphere to provide more useful
services to the world. Supplement 4. Background on existing
space remote sensing systems. A collection of material on
existing operational and research remote sensing satellites and
their associated ground systems plus the relationship to
operational non-NASA data sources. Supplement 5. NEEDS - NASA
end-to-end data system. A system to contain the increasing
amounts of data received from sensors plus a data handling system
to define end-to-end data which will enable real time management.

Technical Support Package for Data Systems and Discipline
Panels - Workshop Document #6 Draft 2 General Discussion
OSTA Alternative Data Systems Concepts
September 11, 1979

Objective of OSTA system is to provide timely,
affordable, accessible, and readily usable multi-source data
products. The concept is a vertically integrated project
efficient for one-source-to-one-user relationships and, a modular
data set which can store data for some period of time to support
many-to-many relationships. Evaluation criteria concepts, pre-
processing/archival systems, on-board systems and relay systems
are all included in the discussion for potential future concepts.

Technical Support Package for Data Systems
OSTA Alternative Data Systems Concepts Workshop Document #6
Revision 1 General Discussion September 20, 1979

Slightly more extensive than workshop Document #6 Draft
2. Includes more discussion on users, pre-processing, and non-
NASA systems discipline oriented information extraction systems.

Instructions for Data Systems Panels
Workshop Document #7 Draft 2 September 11, 1979

Instructions are listed for each systems panel. The
panels are required to identify, critique, prioritize, and
recommend to OSTA the data systems organization, and development
approach. Each panels evaluation and recommendations must
concern its own speciality plus relationships with other
specialities. All panels are required to follow a general
consistent approach in producing their reports. The main
procedure for doing so is outlined in the discussion.

Instructions for Data Systems Panels
Workshop Document #7 Revision 1 September 20, 1979

Revised edition of instructions for each system panel
with the main difference in the revision of the question each
panel must answer in their report.

Detailed Workshop Agenda and Report Outline
Workshop Document #8 September 20, 1979

A detailed agenda of October 9-12, 1979 workshop with
sessions, times, activities, and places all documented. A
detailed format outline for reports of each discipline is
included.

Technical Support Package to Data Systems and Discipline

Panels. Attachments to User Requirements and Data Systems
Survey. Workshop Document #9a. October 4, 1979

Supplement to workshop document #9 and explains primary
data base and traffic volume in mid 80's. It covers all
satellite platforms and sensors according to discipline panels.

Instructions for Discipline Panels
Workshop Document #10 Draft September 20, 1979

Each panel requires evaluations, and recommendations
concerning its own discipline plus relationships with other
disciplines where appropriate. Each panel is to follow a
consistent general approach, which is outlined in this
discussion, to produce their reports.

Initial Statement of Discipline Priority Needs and Key
Elements of Data Systems. Workshop Document #15
October 3, 1979

Recommended approaches, purpose, and implementation of
each discipline panel. The general purpose of the panels are to
identify priority needs of each discipline and to identify key
elements of recommended data systems. Each discipline is
recommended an approach which contains three types of ADS
functions: cataloging, moving data, and spatial services. Also
for each discipline are recommended ADS implementation in three
phases. Phase I is to establish catalogs of NASA electronic
data. Phase II is to augment catalogs by adding non-NASA data,
provide registrations as a special service, and provide
mechanisms for moving data to users. And Phase III is to add in-
situ data and selected special services to include graphs, maps,
or tables.

CSTA Data Systems Planning Workshop - PANEL REPORTS

Technical Support Package for Data Systems and Discipline
Panels. 3 Attachments to User Requirements and Data Systems
Survey. September 20, 1979

1) J. Gitelman. Summaries of Disciplines, Data
Generators, and Users
Contains tables indicating each discipline with
its corresponding data generator, location, auxiliary data banks,
CSTA users, and non-CSTA projects.
2) J. Earle Painter Summary of Primary and Auxiliary
Data Bases by Discipline
A listing of each discipline with its
corresponding primary and auxiliary data bases. This includes
state, local, federal, and individual investigators for each
discipline with a list of the eleven contributing data banks.

o) A. Villasenor Summaries of Data Systems Requirements
Contains tables of operational mission life times
for satellites, estimated data rates for existing and planned
missions, data system facilities, and archive product use for each
discipline.

Draft 2 Report. Part I: Super Summary
ADS Study Office GSFC/100.7 November 11. 1979

To identify and evaluate OSTA data system concepts
which could provide users with readily usable data that is
intended to form a firm decision-making basis for recommending
long-term planning to OSTA.   Also, the need, objectives,
mechanics, requirements, and management recommendations for the
Strawman Data concept and each discipline are included.

Panel Report Summaries and Recommended Overall System Data
Concept. Draft 2 Report. Part II November 11. 1979

Executive summaries of the 11 workshop panels are
reported, plus: an analysis of the discipline reports and
recommendations, a commonality analysis of the data situation and
the recommended Strawman Data Concept. Within each discipline
summary it was realized that NASA acquired space data is only
minimally utilized because of the deficiences in the current data
system.

Land Resources Panel 1b

Report includes a detailed set of requirements,
recommendations, conclusions, and needs. The conclusions include
a need for an overall integrated data system: the system should
be NASA and related remote sensing research/user oriented: the
overall data system should be an end-to-end system; and
subsequent planning for the system must continue to have
significant discipline researcher/user interaction.
Recommendations by the land resource panel include special
emphasis on the development of a cataloging function for data,
electronic ordering of the data, the use of electronic
transmission of the data for the researchers/users, archiving of
the data raw and corrected, and following the data requirements
laid out in table format in the report for spectral, spatial, and
temporal resolution of the data.

Hydrology Panel Report

To better explore the application of remotely sensed
data and perform a research process that combines and integrates
data from many sources the principle research will, first of all,
include universities and their research personnel and federal
research installations supported by private industry.   As
successes are achieved technology transfer efforts should expand
to include state, local, and federal agencies.   A three phase

effort should be undertaken to tie conventional and remote
sensing data for hydrologic studies. These efforts include
digital cataloging of data sources, improving the delivery of
level one data (system corrected and georeferenced) to users, and
improving the reformatting and registration of data from various
sources. A need for an appropriate management philosophy,
management emphasis, and framework with an increased emphasis
within NASA and among others to fully process data gathered is
critical. The project should have strong science/resource
management review and staff, and include necessary team personnel
to define algorithms and couple data sets.

Geology and Geodynamics Panel
Draft 1 Report Panel #3 October 9-12, 1979

Space technology is needed to develop a non-renewable
resource program and to conduct research to increase the
effectiveness of non-renewable resource assessment, exploration
development and global scale management. The goals are to develop
global capabilities for determining the composition and geometry
of materials, and to improve the understanding of the geometry,
composition and evolution of the earth as it relates to non-
renewable resource investigations. Recommendations for the
geology/geodynamic panel include the need for readily accessible
data. NASA space-acquired data should be uniformly gridded, all
NASA data should have a full pedigree with respect to calibrations
and accuracy of measurements, and data delivery time need
improvement.

Dr. David Atlas Report of the Atmosphere
Draft 2 Panel #4 February 21, 1980

Global weather, climate, severe storms, upper
atmosphere, and air quality are reviewed individually with the
main emphasis on exploring and developing capabilities to collect
and digest data on such a dynamic subject. A goal of the
atmospheric panel is to obtain efficient and effective use of a
commonality of observations on a synoptic scale behavior of the
atmosphere. Recommendations include greater accessibility of
data, processing past and present data to a point where it can be
used, preparation of individual data sets, and development of
cataloging system and data directory.

Oceans Panel Report - Preliminary
Workshop Panel #5 October 28, 1979

User requirements consist of 2 elements: the research
element that conducts scientific investigation in ocean
disciplines and includes investigators in academic, government,
and in the private sector; and the commercial element which is
in response to NASA's role as a technology transfer agent, that
carrys out small pilot demonstrations using satellite derived
observations. Also discussed are user requirements including the

need for more ocean data, a comprehensive approach to acquire surface truth as well as satellite data, a partnership with the ocean research community, and the need to have data more available to reduce cost. Recommendations include a more responsive approach to user needs in ocean data concerns and solutions should be forward looking.

Data Acquisition, Distribution and Operations Report Final
Panel #8 December 14, 1979

There are notably more significant management than technical issues which are limiting and will continue to limit the application and use of data communications technology. The most significant issue is a language gap between the statements describing the discipline needs and statements of needs to establish data ransport system. Growth of the ADS system to perform more functions is not likely to run into significant technical limitations in the short run. Limitations exis in the cost of delivering high bandwidth data to remote locations. Data transport systems must have built-in use/cost accounting capabilities. Transport of products that aren't electronically represented should use available technology such as physical transport and electronic transport.

Information Extraction, Processing, and Facilities Panel
Report Final #9 November 19, 1979

Addresses major issues in the construction of OSTA Data Systems and functions of ADS to users. Adopts definitions for levels of processing, determines functions served by ADS, examines management and information extraction alternatives for extraction of information from data calibrated to engineering units. Serves as a baseline for more comprehensive planning in the future.

Applications of Remote Sensing

J.P. Ford, J.B. Cimino, C. Elachi
    Space Shuttle Columbia Views the World With Imaging Radar
    The SIR-A Experiment
    JPL Publication 82-95 January 1, 1983

A selection of 50 scenes displaying the capability of the microwave remote sensor system for geology, hydrology, agriculture, forest cover, ocean surface features, and prominent man-made structures. The images are accompanied by a Landsat MSS or Seasat image of the same scene for comparison. Differences of the scenes are related to imaging systems and to seasonal or other changes within the time interval of acquisition of the images. The images are grouped into main classes that illustrate a wide range of earth surface features:

resources. renewable resources. oceanographic features. and man-
made features.

E.L. Peck, E.R. Johnson, T.N. Keefer
    Creating a Bridge Between Remote Sensing and Hydrologic
    Models. NASA-CR-170517 Final Report January 31, 1980

        Part of a long-term study to determine the suitability
of present and planned remote sensing capabilities for usable
hydrologic models. The need for the study results from low
accuracy of land phase models of the hydrologic variables
measured by different sensors. Study presented is to bridge the
gap that prevents effective use of remotely sensed data by
obtaining a computational algorithm to modify the model to a one-
to-one relationship between remotely sensed variables.

The SIR-B Science Plan
    Imaging Radar Science Working Group
    JPL Publication 82-79 December 1, 1982

        Serves as a reference for the types of geoscientific.
sensor, and processing experiments which are possible and provide
broad guidelines for the organization of the mission. This is a
significant step in understanding the optimum viewing geometries
for various applications. both scientific and technological. that
can be conducted. Includes sensor performance capabilities.
processing. and information extraction.

Proceedings of SPIE Volume 281
    Computer-Assisted Photographic Interpretation Research at
    U.S. Army Engineer Topographic Laboratories (USAETL)
    April 21-23, 1981

        A program in computer assisted photo-interpretation
research has produced a photographic interpreter analyzing high
resolution aerial photography which interfaces directly to a
computer and geographic information system. Initial capabilities
include point positioning. mensuration. stereoscopic area search.
GIS creation and playback. and elevation data extraction. New
capabilities include stereographic superposition. a digital image
work station. and integration of panoramic Optical Bar camera
photography as GIS sources. This project is conceived to be an
evolutionary approach to the digital cartographic feature
extraction program.

California Remote Sensing Committee
    Recommendations to the State Conservationist
    March 31, 1983

        Reviews the current uses of remote sensing. evaluates
the potential for remote sensing application in California.

recommends applications for California programs, and recommend training and equipment which will facilitate the use of remote sensing in California.

E.S. Johnson, E.L. Peck, T.N. Keefer
Combining Remotely Sensed and Other Measurements for Hydrologic Aerial Averages. NASA-CR-170457
October 31, 1982

Describes a method for combining measurements with widely divergent spatial scales and accuracies to develop a sound theoretical base for making model modifications. The scope of the study is limited to the problem of assigning weights to point data, line data, and distributed data for entry into a model.

R.B. MacDonald, A.G. Houston, & R.P. Heydon (NASA/JSC), D.B. Botkin, J.E. Estes, & A.H. Strahler (UCSB)
Reprint for the Seventh International Symposium Machine Processing of Remotely Sensed Data.
Monitoring Global Vegetation June 23-26, 1981

Inventory and monitoring of the world's vegetation on its current status and temporal dynamics is discussed. The need and capability for incorporating satellite imagery, sensors, computers, analysis techniques and ancillary data, together with coordinated implementation approaches, can be effectively tied together, which on a global scale, could aid in the monitoring of the Earth's vegetation.

A Study to Identify Research Issues in the Area of Electromagnetic Measurements and Signal Handling of Remotely Sensed Data.
NASA NAS-9-16390 July 1982

Research recommendations and approaches for handling remotely sensed data are discussed. Fundamental research in minimization of vibrational and thermal sensitivity of instantaneous sensor boresight that will aide in future registration and rectification of images. Research in widefield large aperature optics point spread function analysis and measurement should be supported. Strong support of research in detector array will aide in uniformity and stability of the arrays. Efforts are needed in radiometric calibration at all wavelength ranges for precision and accuracy and when coupled with on-board radiometric corrections would reduce ground data processing loads. Research in active and passive microwave signal handling and processing is recommended for better extraction and registration techniques. An overall philosophy in ground data processing is recommended for optimal placement of data against criteria of cost, time, and accuracy. Support for a total system design and simulation is recommended.

Frederick J. Doyle
    Revised Status of Satellite Remote Sensing Programs
    International Society for Photogrammetry and Remote Sensing
    April 1982

        Major developments of remote sensing programs in the
beginning of 1982 are discussed. The Space Shuttle, Multi-
Mission Modular Spacecraft, Tracking and Data Relay Satellite
System, Thematic Mapper, and The Global Positioning system are
all faced with policy issues effecting the building, operating,
and managing of these systems. Issues include: elimination of
expendable launch vehicles, elimination of ground station
tracking and data network, management of operational earth
observation spacecraft, and commercial operation of remote
sensing satellite systems.

E.L. Peck, T.N. Keefer, & E.R. Johnson: Hydex Corporation
    Strategies for Using Remotely Sensed Data In Hydrologic
    Models.
    NASA-CR-66729 July 31, 1981

        Assess capabilities of current and planned remote
sensing systems for improving commonly used hydrologic and river
forecasting models. Provides the basis for evaluating the
usefulness of remote sensing measurements for hydrologic models
in their present state and with minor modifications.

Albert Rango
    Remote Sensing of Snow and Ice: A Review of Research in the
    U.S. 1975-1978
    NASA Technical Memorandum 79713 Febuary 1979

        Entails research work from 1975-1978 in snowcover
mapping, snowmelt runoff forecasting, demonstration projects,
snow water equivalent and free water content determination,
glaciers, river and lake ice, and sea ice. This research is
investigated through radiative transfer modeling, passive and
active microwave monitoring, and landsat imagery.

Michael J. Abrams
    Computer Image Processing - Geologic Applications
    JPL Publication 78-34   June 1, 1979

        Computer image processing of digital data performed to
support geological studies by relating mineral content to
spectral reflectance, determining the influence of environmental
factors, and improving image processing techniques.

Irrigation Management With Remote Sensing
    Final Technical Report for NASA/Goddard
    SISL-481-81-02 Febuary 1980

Irrigation scheduling information system (ISIS) proposed to provide a gatewaed system for access to a farmer. ISIS with remote sensing is believed to improve irrigation efficiency and monitor water allocation.

John E. Estes, John R. Jensen, & David S. Simonett
    Impacts of Remote Sensing on United States Geography
    March 1980

        Examines the impacts of remote sensing on geography. The unique information provided by remote sensing techniques and its information potential as a powerful tool in the logic reasoning process for data analysis is emphasized. Use of remote sensing is predicted to change perceptions, methods of analysis, models, and paradigms.

Robert N. Colwell
    Coast Watch: Water Quality Mapping of the Entire San Francisco Bay and Delta from Landsat MSS Data
    Cal Space Institute-RS80-1   Series 23 Issue 6
    April 1982

        Landsat MSS data combined with in-situ data for assessment of water quality parameters, salinity, turbidity, suspended solids, and chlorophyll concentration, are used to map water quality in the S.F. and Delta area. Regression models are developed for each water quality parameter from Landsat digital data for 50 sites. The performances of these models were evaluated based on the application of selected models to 20 remaining sites and comparision of observed and simulated values for these sites. The regression models then were extended to the entire study area for mapping of the water quality parameters.

John W. Rouse, Jr.
    Comment on the Effects of Texture on Microwave Emission From Soils - Brief Paper 1980

        Analysis of microwave emission from soils in Arizona, Arkansas, Maryland, and Texas raise doubt that the use of field capacity as the soil moisture parameter will normalize the influence of soil texture.

J.R. Anderson, Ernest Hardy, J.T. Roach, & R.E. Witmer
    A Land Use and Land Cover Classification System for Use with Remote Sensing Data
    Geological Survey Professional Paper 964   USGS Circular 671
    1976

        Classification system that is uniform in categorization at the first and second level and receptive to satellite data.

The system uses features of existing widely used classification which can be derived from remote sensing sources.

Donna J. Peuquet
An Examination of Techniques for Reformatting Digital
Cartographic Data
Part I Raster-to-Vector Process
Part II Vector-to-Raster Process

The majority of techniques in computer-assisted cartography require vector formats. Fast techniques for converting digital data from raster to vector for processing, then vector to raster conversions for plotting. Algorithms are included to perform the conversion processes.

R.E. Arvidson, E.A. Guinness, J.W. Strebeck, G.F. Davies, &
K.J. Schulz
Image Processing Applied to Gravity and Topography Data
Covering the Continental United States
EOS, Vol. 63, No. 18 Pages 261-265 May 4, 1982

Applicability of standard image processing techniques to processing and analyzing large geographic data sets are discussed using image-filtering to produce tone images, digitally registering and overlaying topographic and gravity data, and display of the data to emphasize structural features.

A.B. Kahle, J.P. Schieldge, M.J. Abrams, R.E. Alley, C. Levin
Geologic Application of Thermal Inertia Imaging Using HCMM
Data
JPL Publication 81-55 September 15, 1981

Determination of surface geologic materials based on their thermal properties from HCMM data. Three test sites in the Western U.S. are used. Results of investigations include: partial success of quantitative accurate thermal inertia values from HCMM data due to sensor miscalibrations, radiative transfer in the atmosphere, and varying meterology and elevation measurements; apparent thermal inertia qualitatively represents true thermal inertia; and digital day and night HCMM data allows construction of geologically useful images. Future work will extend these results and examine multivariate use of HCMM with Landsat and Seasat data.

Brent N. Holben, & Christopher O. Justice
The Topographic Effect on Spectral Response from Nadir-
Pointing Sensors
Earth Resources Branch NASA/Goddard September 1980

A model presented to simulate Landsat sensor response was applied to 2 subsets of field data to establish the magnitude

of topographic effect on satellite data. data sets were taken to compare effects of solar elevation and azimuth at different times of the year. 50 pixel values were obtained for high solar elevation data showing that a wide range of pixels can be associated with one cover type due solely to variations in slope angle and aspect.

Bobby G. Junkin
A Method for the Processing and Analysis of Digital Terrain Elevation Data
Earth Resources Laboratory/NASA Report 177 January 1979

A method for the processing and analysis of digital topographic data that can be entered in an interactive data base in the form of slope, slope length, elevation, and aspect angle is presented. Scale factor considerations and mathematical considerations involved in the registration of raw digitized coordinate points to the UTM coordinate system is also discussed.

Micha Pazner
Anti Data Modulation Potential in Expert Systems
A Preliminary Case Study Utilizing Landsat MSS Data
March 1982

Partial anti data modulation effects can be integrated into a software architecture designed to aide a scientist in analyzing an hypothesis. Design considerations for data collection, representation, and processing are aimed at reducing modulation distortions in hypothesis testing.

V.R. Alvarado, Editor
Assessment of the Use of Space Technology in the Monitoring of Oil Spills and Ocean Pollution
Technical Volume NASA/NAS1-15857 April 1980

Applications to the control of ocean pollution where space technology can contribute include communications, ship navigation, remote sensing of pollutants, and data management. The capability to monitor large oceans synoptically through space-borne instruments is feasible and the satisfaction of the user needs is facilitated by coverage from spacecraft, aircraft, anced surface platforms. Information extraction techniques for high certainty, unambiguous detection of oil spills and waste pollution involve the building of a comprehensive baseline of experimental data concerning space and aircraft observations. Development of instruments for broad spatial/temporal coverage, oil quantification methods, techniques for late model input data, and data handling and communication system approaches are investigated.

Application of Remote Sensing to Selected Problems within the

State of California
Space Sciences Laboratory/UCB
Annual Report Series 22 Issue 19 May 1, 1981

The study includes basic research to develop a remote
sensing derivable classification scheme that will enable the
resource manager to solve inventory/management problems. Applied
research to ensure the manager that the classification scheme
provides him with proper information and an applied aspect is
further designed to ensure transfer of technology to the user.

Jeff Dozier
A Method for Satellite Identification of Surface Temperature
Fields of Subpixel Resolution. 1981

A method to identify the magnitudes and the subpixel
areal coverages of two different surface radiant temperatures
from TIROS-N satellite data. Portions of a pixel occupied by
each temperature field, according to the method, assumes that
there are only two temperature fields with a target temperature
and background temperature.

Anne B. Kahle
Derivation of a Thermal Inertia Image from Remotely Sensed
Data. JPL

Thermal inertia of the earths surface is used in
geologic mapping as a complement to surface reflectance data.
Thermal inertia cannot be determined directly but, must be
inferred from radiation temperature measurements over time
combined with a model of surface heating processes. A model
incorporating meterological variables and remotely sensed
temperature data can produce thermal inertia images.

University of California - Berkeley Campus
Remote Sensing of Rice and Related Crops
Remote Sensing Research Program
Space Sciences Laboratory Series 21 Issue 25

Part I - A Four-Decade Overview. Overview of research
to establish the multiband, multidate, and multistage
specifications that appear optimum for the monitoring, yield
prediction, and management of rice in two-test areas: the
Central Valley, CA and Louisana.

Part II - California Case Studies. Case studies of
rice growing areas and rice-related areas. Californias rice
fields can be identified with a high degree of accuracy from
multidate Landsat MSS imagery, supplemented to a limited extent
with aerial photography and ground observation.

Part III - A Detailed Indonesian Case Study.

Development of a crop forecasting technology that applies
specifically to Indonesias annual rice crops. Also, a remote
sensing based system is developed for monitoring wetland rice
production.

Kofi Apenyo
    Survey of Representative Remotely Sensed Data Base Systems
    JPL December 31, 1981

        A synthesis of characteristics of data management
systems relevant to OSTA disciplines that is input into the
derivation of functional requirements.

Armond T. Joyce.   Land Use Applications Requirements Study for
        the ADS Feasibility Study. May 1979

        This is an internal working paper, which defines and
characterizes the data pertinent to land use applications and
summarizes the users and producers of the data.   78
Universities of 27 states are increasing their capabilities to
derive land cover data from Landsat MSS data. The demand for
Landsat MSS data pertains to three main activities: research,
technology transfer, and operational use. The data demand
resembles a bell-shaped curve over time. A simple approach to
estimating future Landsat digital data needs suggests cloud-free
data taken 6 times a year. A prediction of the future demand for
all space-acquired data could not be determined.

John   Estes & Ludwig Eisgruber Information Utilization and
        Evaluation Final Report
        NASA Fundamental Research Program NAS 9-16077

        Identifies issues involving an understanding of remote
sensing technology for data and how data becomes information in
the renewable resources management decision process. Future
predictions show difficult choices among fundamental and applied
research, and development projects that could potentially lead to
improvements in the information systems used to manage renewable
resources. A renewable resource information system must be set
up to improve its remote sensing applications.

The Multispectral Imaging Science Working Group.
    Working Group Reports - Final Report Volume II
    NASA Conference Public 2260 September 1, 1982

        Proceedings of working groups sponsored by NASA, OSSA,
and Earth and Planet Exploration Division. Working group
consists of 4 earth science panels: botany, geography, geology,
and hydrology plus two technical oriented panels concerned with
sensor design and data reduction. Each panel studied and
summarized spectral and spatial characteristics of the sensors

82

surface, specified desired multispectral measuring capabilities and identified gaps in remote sensing for future efforts. Research data needs which come from the 4 disciplines state that higher spatial resolution to address specific problems, finer spectral resolution, exact time of observations, and radiometric accuracy which is dependent on the dynamic ranges of spectral signatures being observed, are all needed for future remote sensing efforts. Use of area array and/or programmable filters can provide spectral flexibility and on-board computation techniques that can be used to select specific spectral bands and spatial resolution. The need to understand the effects of atmosphere and viewing direction on spectral signatures is emphasized.

N84
31747

UNCLAS

D8

Appendix C.
Global Resources Information Systems: A concept...

F.C. Billingsley, J.L. Urena
J.P.L., Pasadena, California

J.E. Estes, J.L. Star
University of California, Santa Barbara

APPENDIX C

Global Resources Information System, Progress Report

John E. Estes, Jeffrey L. Star, Michael J. Cosentino, Lisa J. Mann
University of California, Santa Barbara

NASA Jet Propulsion Laboratory is conducting a study to
define the basic design criteria and operating characteristics of
a Global Resources Informaiton System. As part of this effort
UCSB researchers have been compiling background material and
aiding JPL personnel in this project definition phase of GRIS. A
bibliography of past studies and current work on large scale
information systems has been compiled with partial support from
this grant and is included as appendix B of this document.

All resources, remote sensing and general library, at UCSB
were examined in preparation of this bibliography. This
bibliography presently in draft form will be of real value as we
progress in our research on large scale spatial information
systems. The materail in this bibliography will be continuously
updated throughout the lifetime of this grant.

A draft report "Global Information Sytem: A Concept
Paper (JPL Publication D-1524, F.C. Billingsley, J.L. Urena, J.E.
Estes and J.L. Star) is now being circulated for comments. A copy
of this report is included in Appendix C. of this document.

In the coming months, UCSB and JPL researchers will continue
to interact on this project and aid JPL in defining user
requirements and operational scenarios.

CONTENTS

SECTION I

INTRODUCTION

Future National Aeronautics and Space Administration (NASA) flight missions and, in particular, the planned Space Station will provide an unprecedented potential for advancement in the Earth Sciences in the next decade. However, the new requirements that will be levied by the scientific community will also create a challenge for the implementation of information systems that will improve scientific productivity to a level that allows for the realization of this potential.

In man's attempt to understand the Earth and its environment, enormous amounts of data are being gathered from many different sources. The extreme dispersion and sheer quantity of this data historically have made it extremely difficult for space and terrestrial applications scientists to access data of particular relevance to their work or even to learn of its existence. Problem areas of greatest concern have been the poor quality, poor usability, and inaccessibility of relevant data.

The demands upon NASA's information interface to the user community are evolving along with the research objectives. The complexity of the environmental and resource processes that affect the habitability of the Earth requires a multi-disciplinary approach to the understanding of the natural phenomena and their dynamics. Increasingly, the requirements involve comprehensive, multi-mission, multi-disciplinary investigations that require data from many sources, both from space sensors and from conventional sources. In addition, it has become apparent that certain disciplines interact, using parameters generated within the purview of other disciplines. There are commonalities of needed source data and parameters among the uses and users in many Earth Science discipline areas. One study (Reference 1-1) shows that, considering all disciplines and parameters, on the average over 3 1/2 disciplines are served by each parameter derived, and each parameter could be measured by over ten sensors.

The task of locating and assembling data from several heterogeneous sources, and then integrating the data into usable correlated data sets, has become a technically formidable undertaking. This work frequently consumes a sizable fraction of the total resources available for a research investigation.

The earth-watching satellites of NASA and other national and international agencies provide a wide spectrum of detailed data about the earth's resources. Heretofore, the state of the art in information science was such that, to perform broad-scale scientific research, drawing upon multiple data sets from different sensors and satellites meant long, laborious, expensive periods of data gathering and manipulation. Today, with advances in the technologies of data storage, networking, and data management, it is possible to develop a system within practical costs to allow efficient remote access to multiple distributed data bases and to enable productive global-scale research.

In recognition of this growing need, the House Committee on Science and Technology of the United States Congress directed NASA to prepare a "preliminary program plan for a Global Resources Information System utilizing an interactive network of relevant data bases including program scope, technology, needs, and resource requirements."

In 1978, recognizing the need and the technological opportunity, NASA initiated a coordinated, multidisciplinary activity aimed at making satellite data accessible on-line to the NASA applications research community and providing within the system the processing capabilities necessary to enable global-scale, multidisciplinary research based on multi-sensor data sets. This activity involved the systematic development of improved techniques for data-base management, networking, data standards, and advanced processing systems. The scope of the task is large and complex; therefore, it was structured around a series of distinct science-discipline-oriented "pilot systems" and supportive technology development tasks.

Today, pilot systems have begun to serve the needs of the atmospheres/ climate and oceanographic communities; the Pilot Climate Data System and the Pilot Ocean Data System are now operative and supporting scientific research. This activity continues with the development of two new pilot systems oriented toward the land-related science and planetary science research communities, respectively, together with on-going work with applications disciplines to develop and demonstrate improved tools and techniques for researchers in earth resources, oceanography, atmosphere physics, climate research, and planetary science.

These pilot information systems, together with the supporting technology developments necessary to accomplish the stated goals, can provide the foundation for a Global Resources Information System (GRIS), a broad-based, multidisciplinary information system to facilitate integrated research involving global data sets potentially available from satellite remote sensing and other conventional sources to address global processes involving the land, air, and water.

In its initial stages, the Global Resources Information System will serve NASA and NASA-supported and cooperating scientists involved in applied science research within selected universities and other U.S. institutions and federal agencies. However, the global nature of the GRIS concept and of future flight missions (like the Space Station) that it will be supporting dictates that the scope will have to be broadened increasingly to include scientists and institutions with data bases and research interests in key foreign countries as well.

The range of science disciplines and investigators involved, the number of agencies that will need to cooperate and/or participate, and the international scope of the Global Resources Information System pose a significant challenge. However, it is a challenge that must be accepted if we expect to be prepared to use the full potential that space systems of the next decade represent for the advancement of the sciences of the Earth and the understanding of our planet.

The purpose of this report is to present a preliminary concept for a Global Resources Information System and to propose an initial approach to a Program Plan. The scientific rationale and the scientific objectives of GRIS are discussed in Section II. Sections III and IV present the preliminary GRIS concept and system overview. Section V includes a discussion of some of the technical and political implications of the GRIS concept, and Section V suggests an initial approach to the creation of a Program Plan.

SECTION II

SCIENCE RATIONALE

The science rationale for a Global Resources Information System has at its foundation the concept that the Earth is one large, complex life-support system made up of discrete, interacting parts. The functioning and inter-relations of these parts (lithosphere, hydrosphere, biosphere, and atmosphere) is not clearly understood. In addition, humanity is subtly and surely intro-ducing changes that impact the balance between and within these parts. As mankind's technological potential has expanded, so too has man's inherent responsibility to employ technology wisely. To many scientists involved in research on all aspects of the Earth's global resource base, it has become clear that in this century, the capability to significantly alter this planet's environment on a global scale has finally been achieved. What may be less apparent to these same environmental scientists is that research has also developed new tools and techniques that offer the potential to observe, analyze, and perhaps control the end results of action affecting the environment. To date, humanity's actions regarding the environment have been haphazard: Man has acted first and only later recognized and attempted to assess the consequences of those actions. Organic carbon deposited by natural processes over millions of years has been extracted from the earth in dozens of years. Until recent years man has used energy stored in fossil reserves, depositing vast quantities of carbon in the atmosphere (estimated to be about four billion tons per year) without regard to the consequences on climate.

The effects of man's actions may be subtle. Agricultural practice, favoring cultivation of legumes, enhances the rate at which nitrogen is fixed by natural processes. In combination with nitrogen fixed inadvertently by combustion or deliberately during the manufacture of fertilizers, this nitrogen may give this planet's global life-support system to new domains, with as yet unpredictable consequences for the air, sea, soil, and biota. Humanity's actions may also be direct. The greatest feat of global engineering, conversion of an estimated 10% of total planetary land area from natural vegetation to agriculture, was carried out without regard to the large-scale consequences. Early farmers did not file environmental impact statements. Even if they had done so, their documents probably would have provided little insight into the long-range consequences of agricultural land conversion. The scientific and technical basis for such a global assessment was lacking then, and to a real extent, is still lacking.

This has been in part due to a lack of adequate data of the type, scale, and quality to address scientific issues that are global in nature. In addition, the complex nature of these problems requires that a multidisciplinary approach be taken. In a society in which specialization has been the modus operandi, interdisciplinary research of the type needed to address global problems is rare indeed. It is, then, a basic premise supporting the need for the Global Resources Information System, that these large-scale (global) problems require multidisciplinary research. It also seems that little truly multidisciplinary, large-scale research is occurring.

Reasons for this lack of research are many, varied, and complex. They range from basic technical problems associated with the collection, archiving, networking, and processing of large data volumes to fundamental factors involved in having scientists with diverse backgrounds meaningfully interact to achieve research goals. Yet, it is just this type of research, in which satellite and aircraft remote sensing is supported by surface sampling and backed by effective information systems, that can provide significant benefits today. Some examples of global-scale problems that could benefit from such a multidisciplinary, remote-sensing, total-information science approach are described below. The fundamental characteristics of these problems are that they are global in scale and require a concerted, long-range research effort to achieve significantly improved understanding.

In each of these examples of global problems, global models are required. Currently, global models for the most part are not available to support many lines of scientific inquiry, and, in many cases, those that are available are inadequate. Global models inherently require global data for their solution. Use of global data requires that they be planned for, acquired, then located and assembled. In this global modeling problem, even if the models were formulated, the quality data required to fully support them are not currently available.

1.   Global Carbon Dioxide Cycle

Several large models are required: Global atmospheric circulation, interaction with the ocean, effect of terrestrial vegetation, release from fossil fuels, other sources of natural carbon dioxide, biological productivity, and relation to other global chemical (P,N,S) cycles, to name several. For this discussion, the thrust is whether remote sensing can provide the necesary data to solve these models, and what is necessary for an adequate information system. This will require such varied information as circulation patterns; latitudinal temperature variations; forest distributions by types and rates of change; evaluation of cultural practices as they affect forest product disposition of carbon in various forest regimes; carbon transfer at the air-sea interface as pH, temperature, and sea state vary; ocean precipitation of carbonates; and fossil-fuel practices. The multidisciplinary nature of the problem and wide gamut of data required must be addressed.

2.   Acid Rain and Pollution Dispersion

Information concerning circulation patterns is required. In addition, knowledge will be required of a number of individual parameters: fuel components and their dispersion, depending upon energy practices; man-made sources such as automotive combustion; possible natural sources; effects on soil pH and water pH; efficac of natural buffering; the identification of specific sources by trace components; and the relative importance of S and N. Can remote sensing of trace materials be of assistance? Given the problems of repeat coverage due to orbit constraints, are better orbits possible? How about a synchronous polar orbit repeatedly going across Kansas with trace pollutant sensors? Could laserbased, absorption-line detectors be of use? If one addresses this problem, what would one do specifically; and what remote sensing supported by what type of information system would be needed?

2-2

3.   Biogeochemical Cycle

        The biochemical cycle is the very foundation of our existence. To
paraphrase recent studies: "In the coming years, we will be able to predict
how much food the Earth will produce while most of that potential food is
still 'on the hoof' or in the ground; this will help us to feed the hungry.
We will be able to know in advance how much fresh water, timber, oil, and
metallic minerals we can safely consume - this will help us to harmonize
growth...." (Reference 2-1). Again, a multidisciplinary appproach is required
(If we strip-mine Montana for coal, for example, how soon can the wheat and
cattle culture again become practical?) as well as the use of global models
and a wide variety of data. The concern is with both aquatic and terrestrial
cycles and their vitality with changes in their environment: upwellings;
water temperature and other coastal processes; water circulation; El Nino
effects; any interactions with the global CO2 problem; effects of soil pH;
rainfall (acid or not); hydrology and water-table problems; primary productivity
in the wetlands; leaf chemistry; and the nutrient value in crops and development
of crops to suit environmental conditions, to name only a few. Figure 2-1,
from a recent NASA study on Global Habitability (Reference 2-2), illustrates
the interdisciplinary nature of the problems and suggests the use of the remote
sensing component. To aid in the understanding of the biogeochemical cycle
and the renewable resources problems, how can coordinated (across disciplines)
remote sensing/information systems approach aid in solving the present models
and allow new models, now untenable, to be developed?

4.   Deforestation, Desertification, and Habitability

        Remote sensing has aided in the understanding of the problems of
deforestation, desertification, and habitability. Again, the problem is
multidisciplinary. Some of the factors involved include soil conditions;
general rainfall and temperature conditions; population pressure and cultivation
practices; forest stripping for fuel, wood products, or for other uses of the
land; land suitability for other uses; and changes in rainfall due to changes
in vegetative cover. This problem interacts with and uses many of the same
data as the biogeochemical-cycle problem: the global CO2 and weather and climate
problems. Worldwide vegetative maps can be assembled, showing changes over
time. Worldwide data bases of temperature, rainfall, and other parameters
can be assembled. Many of the problems are largely social. In this context, a
question is: How can the social sciences benefit from remote sensing with an
overall information science context?

5.   Weather and Climate Monitoring and Prediction

        Weather and climate monitoring and prediction have been a major
contribution of remote sensing. There is no need for an apologia for this
use. The emphasis for this discussion is the interaction of studies conducted
in this discipline with the others, for example, the use of global circulation
and precipitation models with the global CO2 and habitability problems. Research
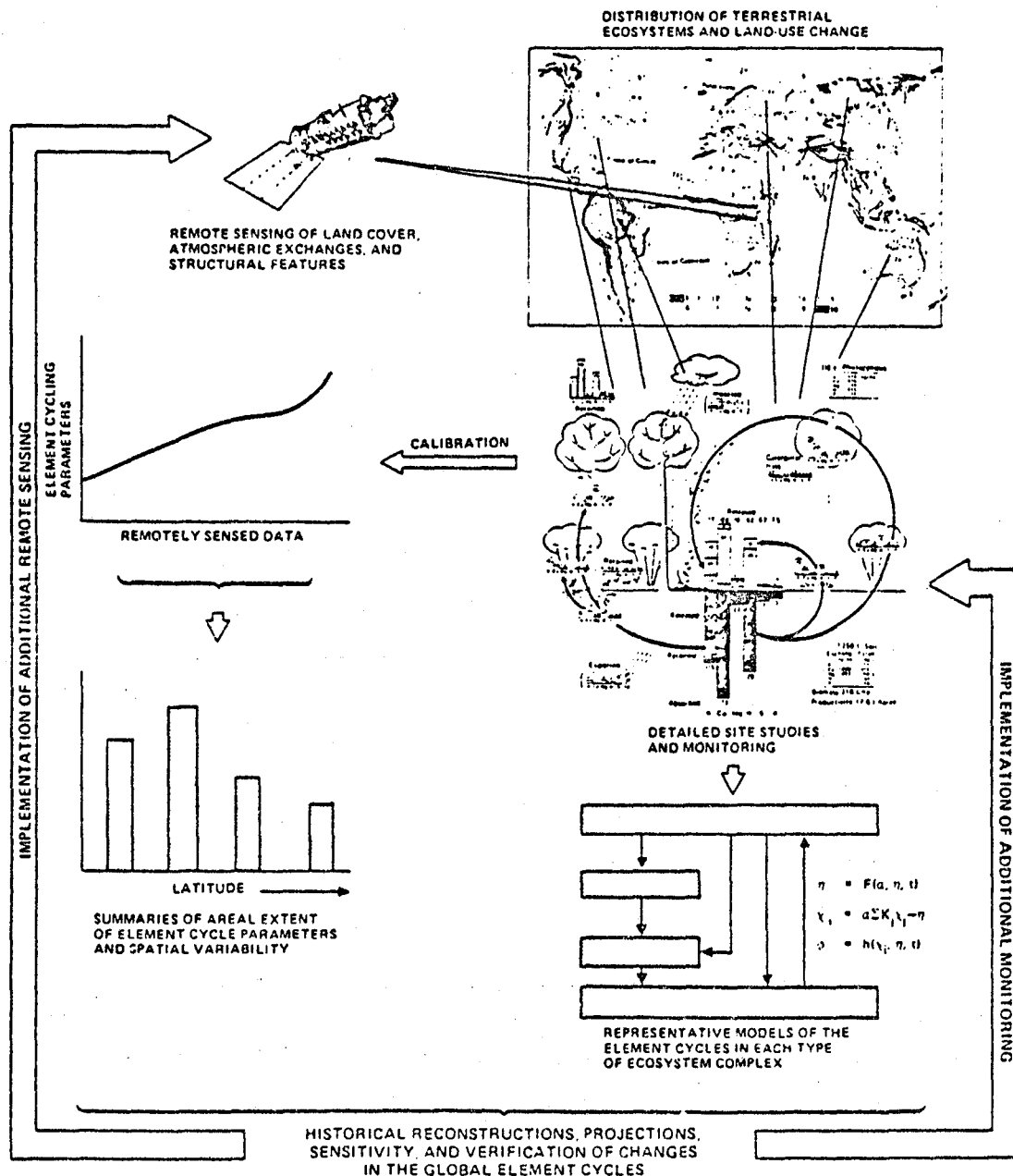
Figure 2-1.   Relationship of Biogeochemical Cycling Research Areas
(Source: see Reference 2-2)

2-4

has shown the potential of correcting Landsat images by flagging areas of high cirrus clouds, which distort normal spectral responses. Can the data distribution techniques developed for this discipline be used to retain more of the old weather data for historical studies?

Other major problems could be discussed, such as: global navigation, global sea conditions and the effects of shipping, ice processes, global topographic mapping, physical oceanography, basic geodesy. All have been, and will be, aided by remote sensing and advanced information systems.

The need for global research is vital. The potential provided by remote sensing and information systems is high. Indeed, when thoughtfully analyzed, the two technologies are inextricably linked. Remote sensing data must reach a user in an appropriate form and time frame. It often achieves its maximum effectiveness when combined with other data in spatially referenced (geographic) information systems. Such information systems achieve their maximum effectiveness when the data they contain are accurate and current. The linkage of these two technologies has not occurred because of fundamental incompatibility. It is a matter of necessity if these global issues are to be realistically addressed.

SECTION III

PRELIMINARY SYSTEM CONCEPT

The prime motivation for the GRIS is the establishment of an effective data-management system to meet the data-access requirements of the scientific community in its multidisciplinary and multi-mission investigations in the Earth Sciences. Such investigations, commonly known as correlative research, use data from many sources for a wide range of applications in a variety of disciplines.

In addition, every discipline can use data common to other disciplines. For example, the climatologist is interested in hydrology, land cover, the upper atmosphere, solar physics, and indeed, in the behavior of other planetary atmospheres as well as the behavior of the Earth's atmosphere and oceans. Each discipline generally must make some of its observations on its own, yet each should not have to make all its observations on its own.

Vast amounts of data on the Earth and its environment are being collected in centers throughout the nation and around the world. However, investigators are, for a variety of reasons, often unable to use more than a small fraction of the existing relevant data. The potential data consumer (investigator) must determine:

(1)   What data collections exist that have data relevant to the investigation?

(2)   Which specific data sets within the relevant collection are pertinent?

(3)   How can the specific data sets be obtained?

(4)   What must be done to convert the data into a useful format?

(5)   What support services and background information exist that would facilitate multi-user, multidisciplinary use of the data?

Because each of the several data sources used in an investigation may have its own access protocols and product forms, the overall process of finding out what data are available, obtaining the data, and integrating the data into a compatible set is currently a technically formidable task. Frequently, this task results in the development of extensive application-unique data access and conversion capabilities.

The GRIS concept is a strategy for lowering the information exchange barriers that exist between participating producers of data and information, and the existing and potential users of that data and information.

The GRIS program will seek to provide key elements to support correlative research in the land, atmosphere, climatic, and oceanographic sciences. Using standard interfaces and operating guidelines, diverse data systems can be integrated to provide the capabilities to access and process multiple geographically dispersed data sets and to develop the necessary procedures and algorithms to derive global resource information.

To accomplish these objectives, GRIS, as currently envisioned, must be capable of performing the following functions:

(1)   Interface to a geographically dispersed, heterogeneous user community via a common-carrier, electronic communications network.

(2)   Provide access to data bases that contain both space and terrestrial data in varying formats and that are managed by dissimilar data management systems.

(3)   Provide substantial ease and uniformity of user access to the data bases.

(4)   Provide extensive assistance in the location of relevant data.

(5)   Enable data to be transferred in a timely and efficient manner.

(6)   Allow concurrent access to the service by multiple remote and local interactive users.

(7)   Allow full control by participating organizations over access to and utilization of their computing and data resources.

(8)   Minimize time, cost, and operational impact of adding (or removing) users or data centers to the GRIS.

(9)   Develop and maintain data and communication standards.

(10)   Support on-line, cataloged processes to perform special functions, allowing controlled computing resource sharing when appropriate.

(11)   Perform special services, such as data integration or active archiving of data, when such services are desired.

(11)   Perform the administrative and management functions necessary for the coordination of the information network.

An important feature is that the data accessed via GRIS will remain under the management and control of the organizations and data centers in which it resides. No data will be managed directly by the GRIS, with the exception of catalogs and directories of information about available data bases and participating organizations and centers and their resources, unless a specific archive is placed under GRIS management for special purposes.

C-2

Discipline-oriented data will reside in discipline-oriented systems. These discipline-oriented systems are based on the concept of a Scientific Data Management Unit (SDMU) (Reference 3-1). An SDMU is an organizational entity responsible for developing and maintaining a "live" data archive in support of a specific scientific objective. A live archive contains data that are in active use, together with a knowledgeable support staff. The staff makes decisions about the structure, organization, and content of the data base required to support the discipline investigations.

The SDMU is typically associated with a central-processing and analysis facility that is a principal user of the data. Other remote users are interconnected through the network systems.

SDMUs and associated computing facilities are created wherever there exists a critical organizational mass with deliverable research objectives.

Examples of SDMUs include NASA's Pilot Oceans Data System (PODS), Pilot Climate Data System (PCDS), and the proposed Pilot Lands Data System and Pilot Planetary Data System, presently under development.

Figure 3-1 is an overview of the GRIS concept. It includes the following elements:

(1)   Archives to retain usable data sets and make them readily available.

(2)   Discipline-oriented information extraction and processing systems to prepare parameter data sets, operate R & D models, develop user-oriented products and displays, and perform scientific analysis.

(3)   A common data cataloging and dissemination network service to allow users to locate, order, access, exchange, and integrate data quickly and at low cost. This would supplement rather than replace the current mailing of tapes and disks.

Section VI of this report presents a more detailed description of some of the features of GRIS, along with a discussion and an assessment of the different networking and information system technologies that are germane.

3-3
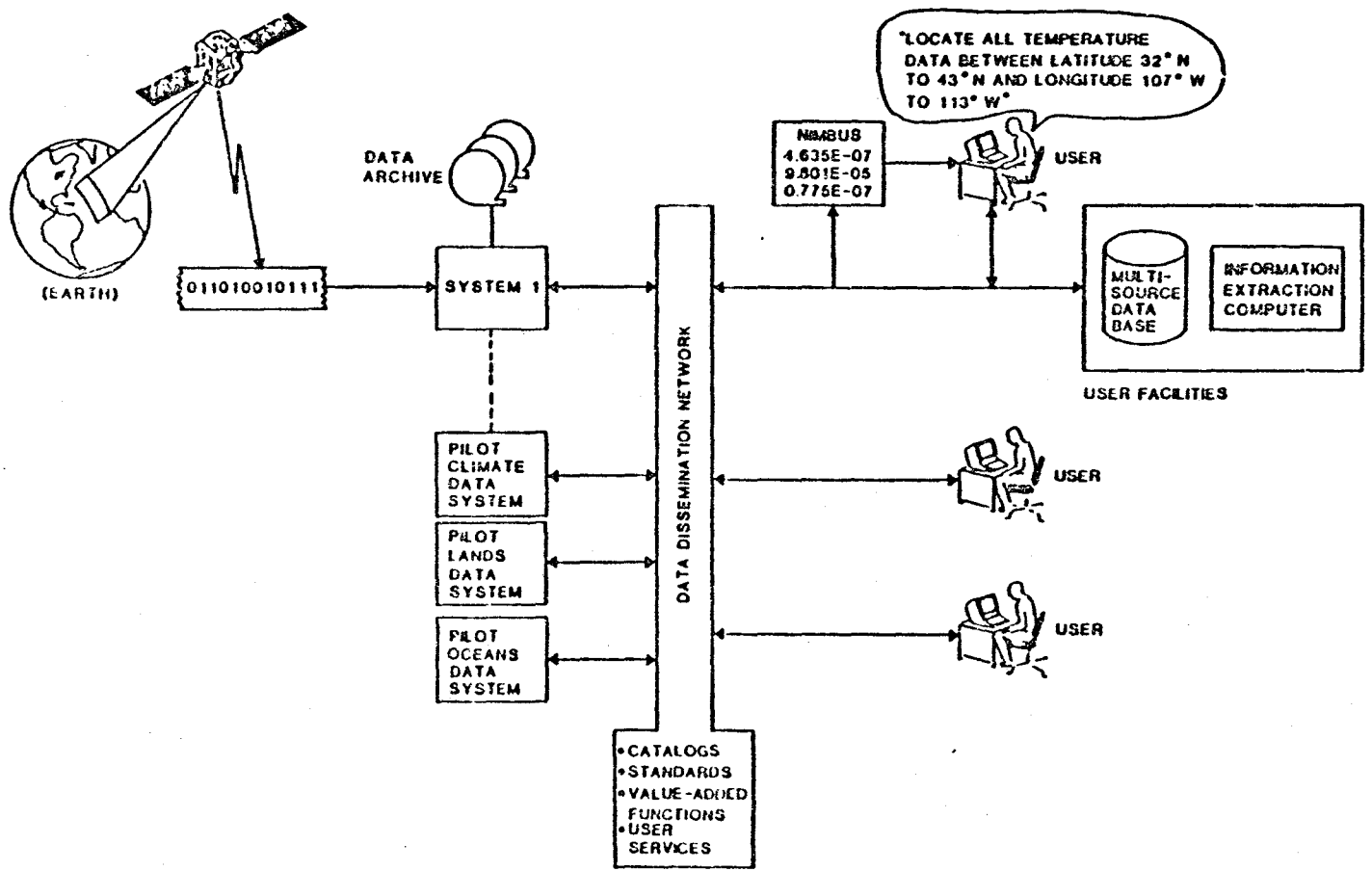
Figure 3-1.   Global Resources Information System (GRIS) Overview

SECTION IV

SYSTEM OVERVIEW

A.   GLOBAL RESOURCES INFORMATION SYSTEM DRIVERS

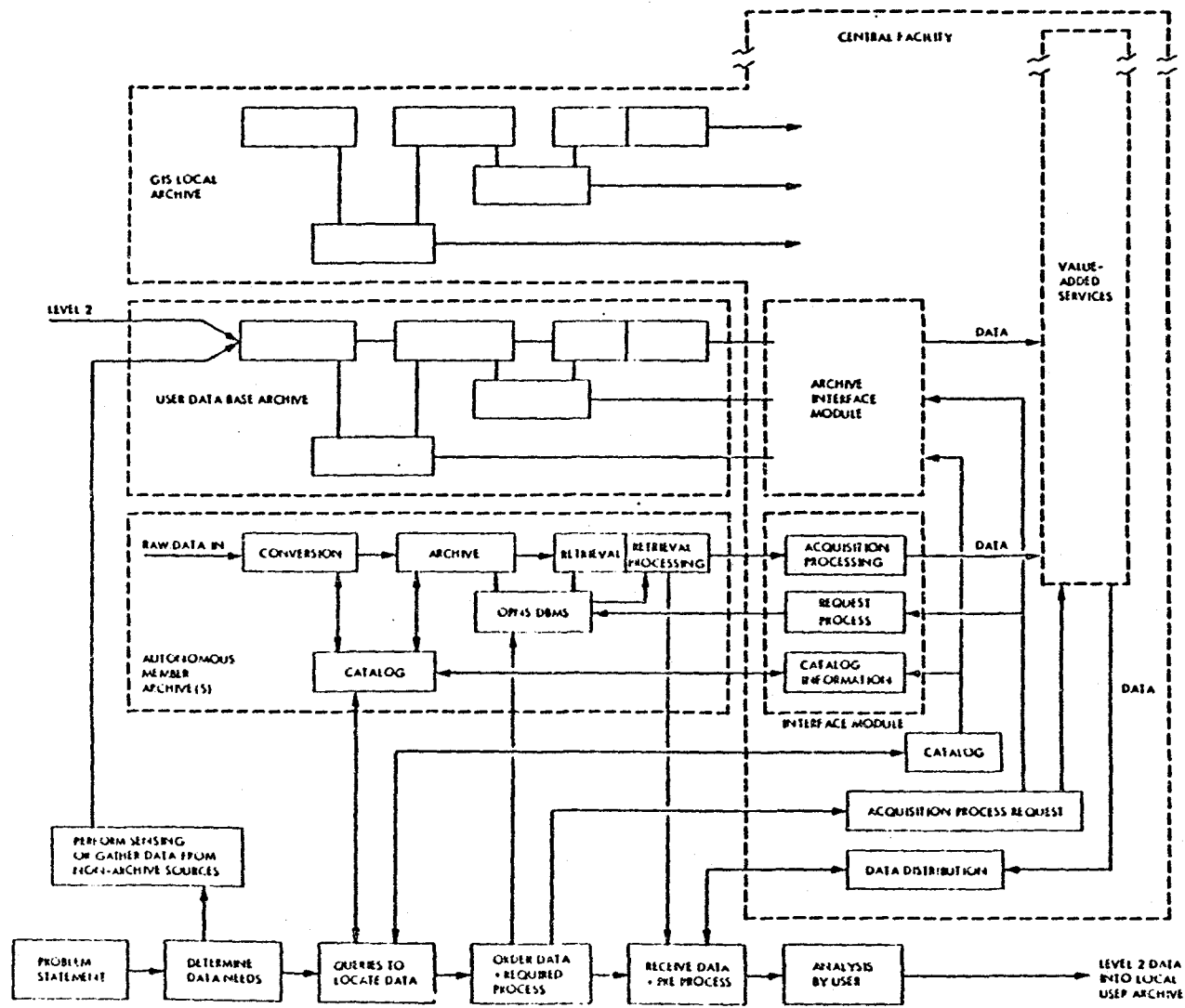A set of generic data handling and analysis requirements results frcm consideration of the large-scale science problems that will be characteristic of future global remote sensing such as will occur from, or in association with, the Space Station:

(1)   Ability to locate required data sets in the various catalogs located in the United States and other countries.

(2)   Ability to move data sets rapidly from the archives to locations where the research is being conducted.

(3)   Ability to register, calibrate, and modify data sets to standards rapidly with minimal manual intervention.

(4)   Ability to perform processes on the group of assembled data sets in near real time, both locally and remotely.

(5)   Ability to communicate research data and technical information between scientists locally and remotely in near real time.

B.   GRIS GENERAL SYSTEM ARCHITECTURE

Integrated software packages known as Geographic Information Systems (GIS) have been developed that can perform all phases of storage, maintenance, retrieval, and analysis of spatial data.   The function that sets them apart from data-base management systems is the ability to perform various degrees of analysis, at the users direction. Accordingly, this analysis will be incorporated within the definition of the GRIS. The general functions of a GIS, of whatever size, are described below.

A GRIS may be thought of as a super GIS, serving many users and accessing many data bases in a systematic, standardized manner. A potential data query and location scenario and the corresponding flow/block diagram will serve as a basis for discussion of factors that relate to the use of geocoded data in such a system (Figure 4-1). Note first that the GRIS will not take over the world or even the established archives. Rather, it will interface with these and provide the coordinating and standardizing so sorely needed. It will have no control over these archives but must understand their structure and cataloging to provide the necessary interfaces. Its structure will be much like that of the telephone service: A centralized organization must be present to organize, provide standards, implement the network and interfaces, perform research to provide future improvements, and provide specialized services. But, just as a telephone customer directly telephones his party without visibly invoking all of the company structure, the GRIS user will be provided with facilitating services as well as specialized ones. Thus, a distributed heterogeneous set

CENTRAL FACILITY

GIS LOCAL
ARCHIVE

VALUE-
ADDED
SERVICES

LEVEL 2

USER DATA BASE ARCHIVE

ARCHIVE
INTERFACE
MODULE

DATA

RAW DATA IN | CONVERSION | ARCHIVE | RETRIEVAL | RETRIEVAL PROCESSING | ACQUISITION PROCESSING | DATA

OPTIS DBMS

REQUEST PROCESS

AUTONOMOUS
MEMBER
ARCHIVE(S)

CATALOG

CATALOG INFORMATION

INTERFACE MODULE

CATALOG

DATA

ACQUISITION PROCESS REQUEST

PERFORM SENSING
OR GATHER DATA FROM
NON-ARCHIVE SOURCES

DATA DISTRIBUTION

PROBLEM STATEMENT | DETERMINE DATA NEEDS | QUERIES TO LOCATE DATA | ORDER DATA + REQUIRED PROCESS | RECEIVE DATA + PRE PROCESS | ANALYSIS BY USER | LEVEL 2 DATA INTO LOCAL USER ARCHIVE

Figure 4-1. A Distributed Geographic Information System

4-2

of data bases is turned into an information system. At the same time, however, the need for timely responses is evident; this may require that the GRIS itself become an archive available for rapid turnaround.

A walkthrough of the diagram will be useful. A user will use his forward model to determine what data is needed to answer a given research problem (lower left). He then queries the catalogs of the various known archives to locate useful data sets. He would also query the GRIS Central catalog, which in turn has access to archive catalogs. The located data are then ordered via either the GIS or the Member Archive operations Data Base Management System (DBMS), which would be the "order desk." Accompanying the order would be requests for the desired retrieval processing. At the same time, value-added services would be requested. On receipt of the data from the GRIS he would proceed with his own program. The GRIS may serve as a temporary or permanent archive, as might the user. Either would maintain archival data compatible with the professional Member Archives. In this manner, all users will have access to all established archives.

In this scenario, the established archives would not change their data structures and would only have to establish suitable catalog interfaces. The archive interface modules of the GIS would perform the necessary standardizing. Their location would be with each archive, to provide a consistent interface to the network.

C.   DATA ARCHIVE PRINCIPLES

To make such a system practical, certain principles should be followed in designing an archive and its data structure:

(1)   To avoid aggregation/disaggregation problems, the data should be stored at the resolution commensurate with its data content, usually at the resolution at which it was obtained.

(2)   If practical, remove the intra-image distortions so that only affine low-order corrections will later be required.

(3)   Store with the data the precision information required to register the data to a well understood reference during retrieval, such as latitude/longitude (i.e., georeference the data). This also provides the required reference information for the data location queries by geographic coordinates.

(4)   Store with the data all relevant ancillary information, such as calibrations, sensor data, processing history.

(5)   During data retrieval, geocode the data. That is, reproject and rescale the data to the grid requested by the analyst, and provide the ancillary data.

(6)   Supply the data in a standard format independent of the incoming data type or sensor, with suitable annotation to allow the analyst to use the various types interchangeably.

D.   SYSTEM MAJOR COMPONENTS

The GRIS consists of the following major components:

1.   Archive Nodes

For this discussion, an archive is a data repository that potentially can be connected to the system for electronic catalog conversations and potential data transfer. The initial supposition is that the data will be required from a number of currently operating archives, that future archives will be established for future NASA missions, that some archives such as the National Space Science Data Center (NSSDC) contain somewhat specialized data, and that other archives such as Eros Data Center (EDC) contain data of more general interest. As these archives are selected to be included, individual decisions must be made concerning the interface modes. The general requirements for an archive node will be to service the catalog searches and requests, to locate and prepare for distribution, and to distribute the requested data. The data reformatting is proposed at the archive nodes to allow all transmissions to be in common style, perhaps using the Standard Format Data Units (SFDUs) being developed. Eventually it will be desirable to maintain an on-line data browse.

2.   The Network

The network would consist of all of the data transport mechanisms as appropriate: mail, leased line, dedicated line, arrangements for Domsat, etc. It could be built on networks such as the NASA Program Support Communications (PSC) facilities, and could contain Local Area Networks to connect users to the Long-Line Network. It would interface to each node with specialized hardware and service both text communications (such as the catalog conversations) and data transfer. As the capabilities can be increased, data transfer progression might be: tabular and low volume point or polygon data, then browse images, then larger images. Interactive capability would initially only be the catalog conversations, then expand to catalog browse, data browse, and (much later) to interactive data analysis.

3.   Support Services Administration

This component would consist of the directory/catalog/dictionary services required to access the various data bases required. It would set the standards for this service and provide catalog query translations, access to the various catalogs, and a supercatalog function. It would also provide the alphanumeric terminals to the member nodes and a search function to locate and connect to new data bases as new data are requested. In short, it is the nerve center of the Pilot: It will provide the administrative and accounting services for the system, including accounting, billing, and standards maintenance. It will include general user support functions such as the administrative user interfaces to the system and the general operations functions required to keep the system running. During implementation, this function would also make the necessary interagency and international arrangements.

4.   System Central Services

        The System Central Services are those data preprocessing services
implemented by the system, as requested by the Steering Committee. This function
will provide value-added services of selected kinds for system and non-system
data upon request from the users. It will also build the capability to perform
value-added services of an algorithmic kind. This service provides the
registration, rectification, mosaicking, media conversions (such as map digitizing),
data format conversions, and related operations. It will be the data outlet
for the system for those data sets not transmitted directly from the archives
to the users and it will maintain an archive of data that it has processed to
avoid the necessity of reprocessing. It will build these capabilities on the
existing capabilities at the various centers and will provide the interface
requirements and specifications for the data terminals.

5.   User Interface Nodes

        User interface nodes are where the research tasks are performed.
They contain the interfaces between the network and the local equipment. The
general concept will also include the local GIS, analysis hardware and software,
and the staff expertise.   This is necessary to fulfill the desire for a
distributed system, not just a networked data base. Network-interfacing
equipment would be under the control of the network function, but all other
functions would be under local control to encourage local developments.

        Compatible, modular hardware will be required at all node locations for
data analysis. This hardware must interface with the terminal hardware and
must support a local GIS.   The data analysis developments being performed at
the various NASA centers and at participating universities will need to be
coordinated and arrangements made to modularize, package, and document for
distribution those algorithms and procedures that are deemed to be generally
useful. All centers are developing or have in place hardware and software
systems; these developments must continue and the available capabilities used
where practical.

SECTION V

ISSUES INVOLVING GRIS IMPLEMENTATION

The scope and the diversity (in terms of users, producers, and resources) involved in the implementation of a Global Resources Information System is the origin of a series of technical and political/organizational implications that go beyond those encountered in the implementation of the existing or planned discipline pilot information systems. The range of science disciplines and investigators involved, the number of agencies and institutions that must be called to participate, and the global scope of GRIS pose a number of technical, managerial, legal, and political questions that require a careful analysis and pose a significant challenge. This challenge must be accepted if scientific productivity is to improve to a level that allows for the realization of the potential that space systems represent during the next decade.

The success of GRIS will depend not only on the correct solution of the technical, technological, and design problems that will be encountered during its implementation but also, to a large extent, on the correct choice of strategies in the areas of management and in the resolution of issues arising from the complexity of the interactions between people and institutions that will be involved.
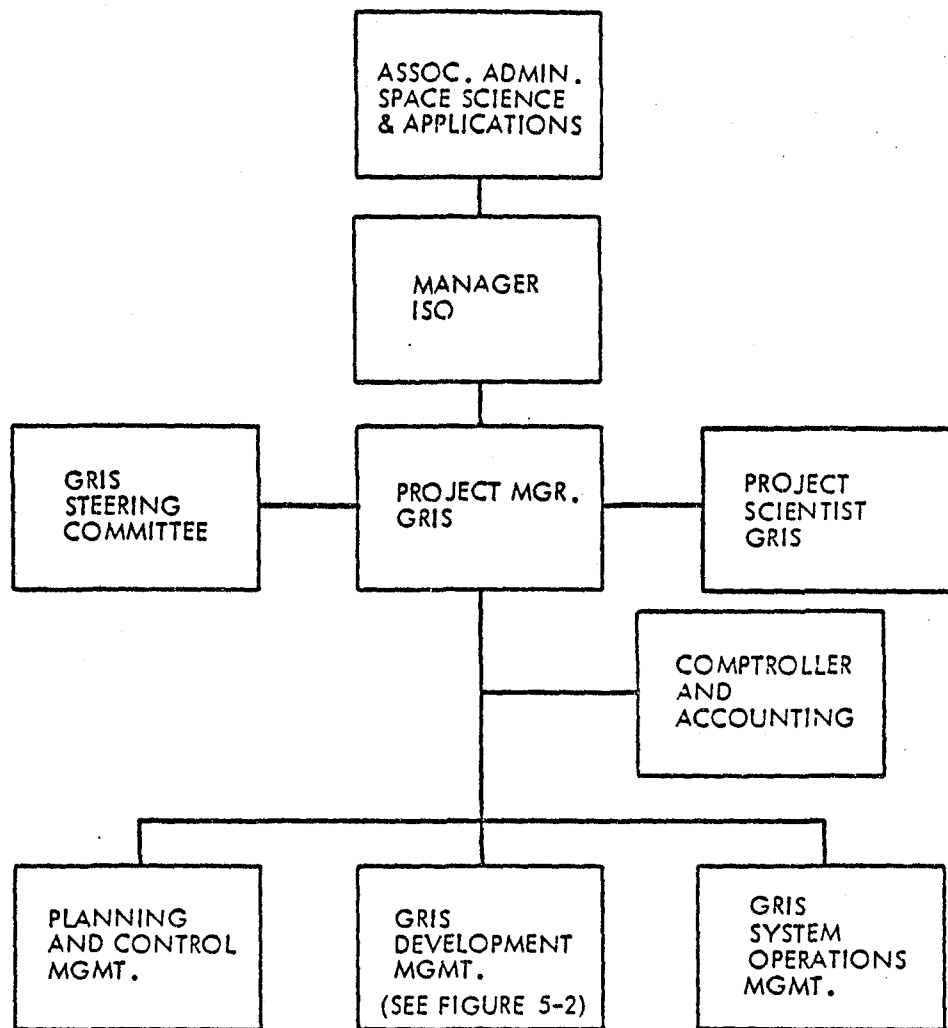
The purpose of this section is to initiate an analysis to identify the principal issues that will affect the GRIS implementation and to provide recommendations where appropiate. Three general areas must be considered:

        (1) Managerial.
        (2) Inter-institutional relationships and coordination.
        (3) Technical.

A.   Managerial

The complexity of the relationships between people and institutions and the number of separate activities that are inherent to the development of GRIS require the development of a strong and well-defined management infrastructure. At the same time, a strong commitment will be needed from NASA Headquarters, involving the Information Systems Office and the Earth Science and Applications Division, to support the development of GRIS through its life cycle. Agreements between appropriate NASA Headquarters offices should be established to cooperate in the planning with future flight missions of the evolutionary growth of GRIS, especially with the Space Station, which is expected to be a main focus of interest to the Earth Sciences research community in the future.

Continued involvement of science users throughout the program life cycle is essential. The GRIS management structure must contain elements in which the scientific-user community is represented to a level that ensures that the program is truly responsive to the real needs of the research community and their scientific objectives. Figures 5-1, 5-2, and 5-3 show a preliminary GRIS organization structure that reflects these principles.

```
┌─────────────────────┐
│ ASSOC. ADMIN.       │
│ SPACE SCIENCE       │
│ & APPLICATIONS      │
└─────────────────────┘
          │
┌─────────────────────┐
│ MANAGER             │
│ ISO                 │
└─────────────────────┘
          │
┌──────────────┐  ┌─────────────────┐  ┌──────────────┐
│ GRIS         │  │ PROJECT MGR.    │  │ PROJECT      │
│ STEERING     │──│ GRIS            │──│ SCIENTIST    │
│ COMMITTEE    │  │                 │  │ GRIS         │
└──────────────┘  └─────────────────┘  └──────────────┘
                       │        ┌─────────────────┐
                       │────────│ COMPTROLLER     │
                       │        │ AND             │
                       │        │ ACCOUNTING      │
                       │        └─────────────────┘
```

- PLANNING AND CONTROL MGMT.
- GRIS DEVELOPMENT MGMT. (SEE FIGURE 5-2)
- GRIS SYSTEM OPERATIONS MGMT.

PLANNING AND CONTROL MGMT.

• PROGRAM PLAN

• SYSTEM DEFINITION & EVALUATION

• SCIENCE DATA PREPARATION

• COORDINATION OF PILOT SYSTEMS

• BUDGET PLANNING

GRIS SYSTEM OPERATIONS MGMT.

• SYSTEM OPERATION & MAINTENANCE

• STANDARDS MAINTENANCE

• TECHNOLOGY SURVEY & EVALUATION

Figure 5-1. Organizational Structure for the Global Resources
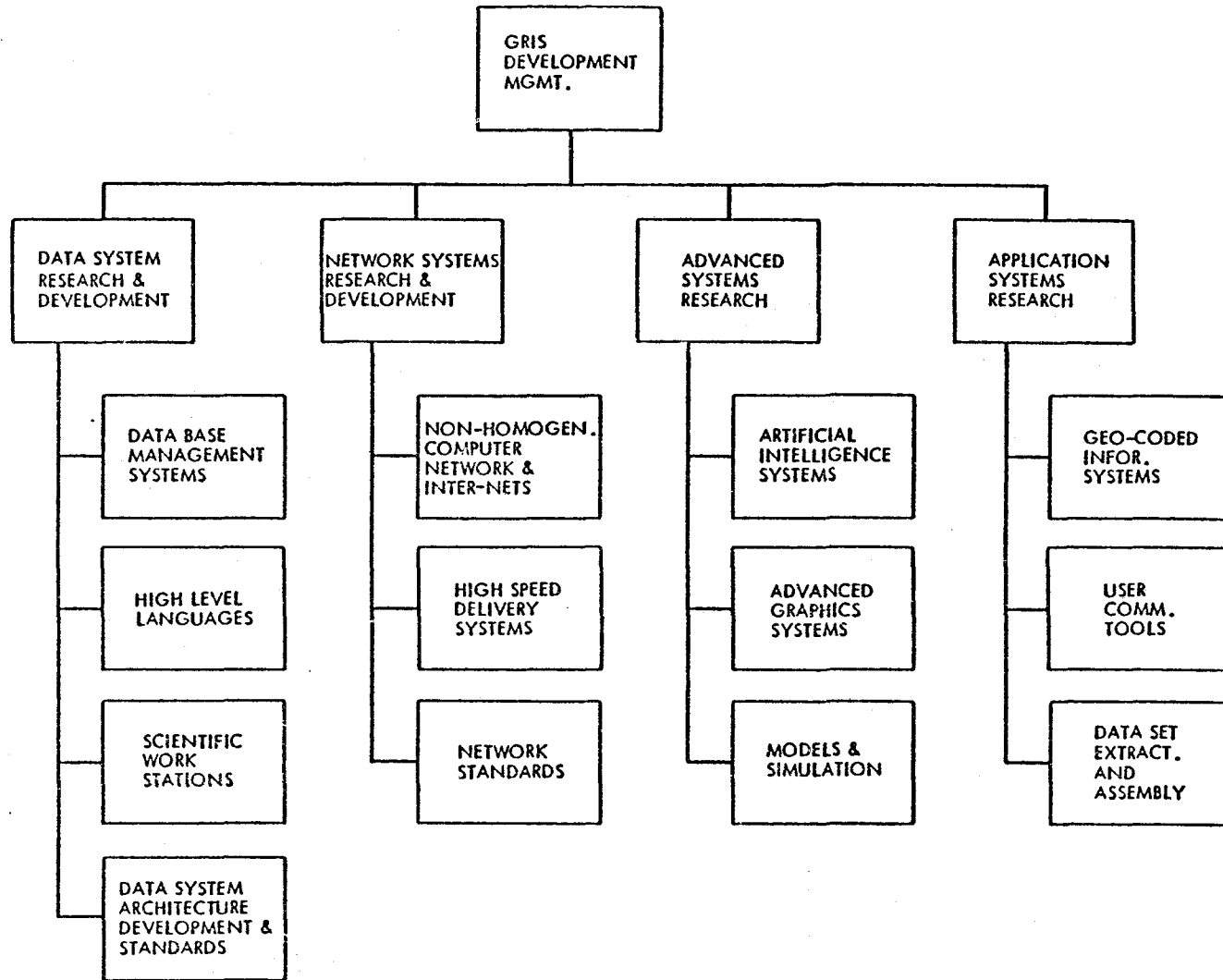   Information System

5-2

Figure 5-2. Development Work Breakdown Structure for the
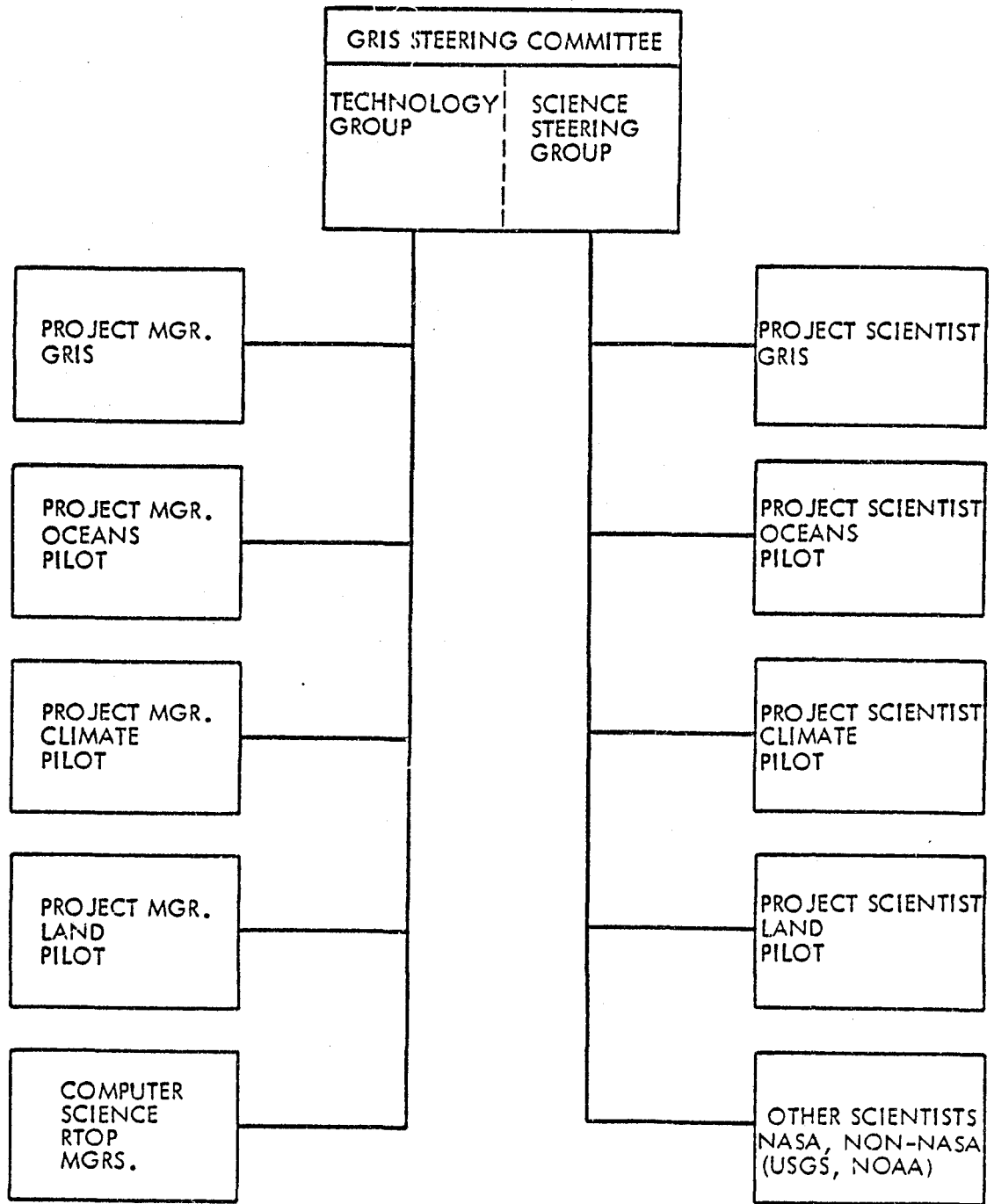Global Resources Information System

Figure 5-3. Steering Committee for the Global Resources Information System

The GRIS Steering Committe (see Figure 5-3) will be composed of a Technology Group and a Science Steering Group. The Science Steering Group will formulate the program's objectives and set forth the requirements for system capabilities in terms of data access, processing, display, and user support.

The Technology Group will identify the relevant information system technologies required and provide guidance in the development and implementation of the system.

The GRIS Project Scientist (see Figure 5-1) will be responsible for providing advice and recommendations for the continued scientific relevance of the program on a continuing basis.

B.   Inter-Institutional Relationships and Coordination

The number and variety of agencies and institutions that are repositories of data of potential interest to the Earth Science research community will require a well-defined set of agreements for cooperation between the involved parties. The sharing and exchange of data and information in a distributed environment raises a series of legal and policy-related questions about the property of data sets and the boundaries of responsibility of the involved organizations.

The issue of irter-institutional and inter-organizational coordination can be viewed from the perspective of the following three levels:

(1)   Among NASA Headquarters offices: The roles and the boundaries of responsibility of the Information System Uffice and the science program offices of the Earth Science and Applications Division will have to be clearly defined through agreements.

(2)   With external agencies and institutions: Federal agencies, such as the USGS and NUAA; and other research institutions, such as universities.

(3)   At the international level: With other foreign agencies and institutions.

The coordination among the participating institutions can be accomplished through a process involving:

(1)   Inter-agency agreements regarding the exchange of data and information as well as the roles and responsibilities associated with their participation in the program.

(2)   Technical and scientific working groups to define the scientific requirements and technical approaches to achieving data and information sharing and exchange.

C.   Technical

The interconnection between diverse, heterogeneous information systems and the number and size of potentially relevant data bases and their geo-referenced nature poses a significant challenge. System design must support a phased, modular-implementation approach so that the design can be validated by building and verifying portions of the system prior to committing to full implementation.

To minimize risks, GRIS development should be based on the integration and testing of proven technologies . However, GRIS should provide a mechanism for close cooperation with the computer science research community for the development of advanced technologies that would eventually be integrated into GRIS. Figure 5-2 shows a preliminary Work Breakdown Structure for the technical activities associated with the development of GRIS.

One of the barriers that exists to computer-based information exchange between producers of space and non-space data and information and the scientific community that makes use of the data and information is that many of the data producers and users have developed unique methods of data handling and processing. These methods include a wide diversity of data descriptions, data formats, and data-exchange protocols.

A GRIS Standards and Guidelines Program should be initiated to facilitate the exchange of technical data and information among the data systems that support the scientific-user community. It would foster common or compatible data formats for multi-source investigations, thereby reducing the number of unique interface definitions and requirements for custom software development by users.

External agencies, such as the National Oceanographic and Atmospheric Administration (NOAA) and the United States Geological Survey (USGS) are large producers of data used by the scientific community involved in Earth Sciences research. It is essential that they are included in the data standards development and participate in the evaluation of the standards implemented under the GRIS auspices. Additionally, some inter-agency standardization efforts, such as the Consultative Committee for Space Data Systems and other organizations, have been developing standards for data and telecommunication systems. Coordination with these organizations is essential to avoid duplication of efforts and to facilitate compatibility with external systems.

Although NASA Headquarters should make all determinations of policy for the program, the development of new standards and the handling of requests for modification of existing standards will be under the control of a Standards Control Board, which will have ultimate responsibility over the operations related to the Standards and Guidelines Program.

SECTION VI

PROPOSED APPROACH

The underlying motivation of GRIS is to establish an effective and efficient information management system to meet the data access requirements of NASA and NASA-related scientists conducting large-scale, multidisciplinary, multi-mission scientific investigations. The objective of GRIS is to develop an interactive data dissemination network that will provide research access to data on a global scale to serve the needs of the atmosphere/climate, oceanographic, and land-sciences communities. An initial step toward accomplishing this task is to examine the potential for linking existing and planned NASA pilot data systems. Within the context of GRIS, this linkage would facilitate research of a type and on a scale not currently planned under the Pilot Climate Data System, Pilot Ocean Data Systems, and Land Data Systems.

Figure 3-1 presents a potential structure for a GRIS. To the extent that these existing and planned pilots facilitate multidisciplinary research, they will support GRIS objectives. To the extent that these existing and planned pilots support global research, they will support GRIS. Yet GRIS must go farther: GRIS should support access to international data banks and facilitate linkages between scientists from many disciplines in a manner not found in the existing pilots.

GRIS should, therefore, undertake a study to determine an effective methods for linking existing institutions. Such study would be most appropriate at this time. Efficiency may be achieved by identifying centers and institutions with existing or proposed links to more than one Pilot Data System. This type of study could examine the best way to effect linkages minimizing redundance and ensuring efficient access to multidisciplinary use of pilot facilities.

To implement GRIS effectively we must also begin our examination of the international implications implicit in the development of such a system. Questions we must ask include: What international data banks must be accessed? Are there restrictions on their access? What international organizations and/or institutions may be involved? What is the most appropriate way to provide multidisciplinary, multinational research team access to and processing capabilities for NASA and related data sets? What problems will hinder this access? We must begin to anticipate these and other issues if we truly hope to achieve a Global Resources Information System capability.

Finally, to test the concept of GRIS a limited number of success-oriented, multidisciplinary research projects, linking ocean, land, and atmospheric scientists within the United States should be initiated. Researchers conducting these studies would be strongly encouraged to cooperate with international organizations yet need not necessarily initially include such individuals as co-investigators. These studies would be initiated only after careful study of the scientific justification and international implications of the proposed projects. These projects would be used to test the feasibility and viability of the GRIS concept. Projects selected should have scientifically significant, specific goals and objectives.

Realistic costs and schedules should also be developed. The premise
is that successful demonstration of the GRIS concept will generate enthusiasm,
and this enthusiasm will in turn provide further impetus for GRIS implementation.
Figure 6-1 suggests a phasing for the studies outlined above. It should be
emphasized that, while development in each of these areas could be undertaken
independently, successful implementation of the GRIS concept will require that
work be done in each area. Therefore, a phase-integrated approach would be
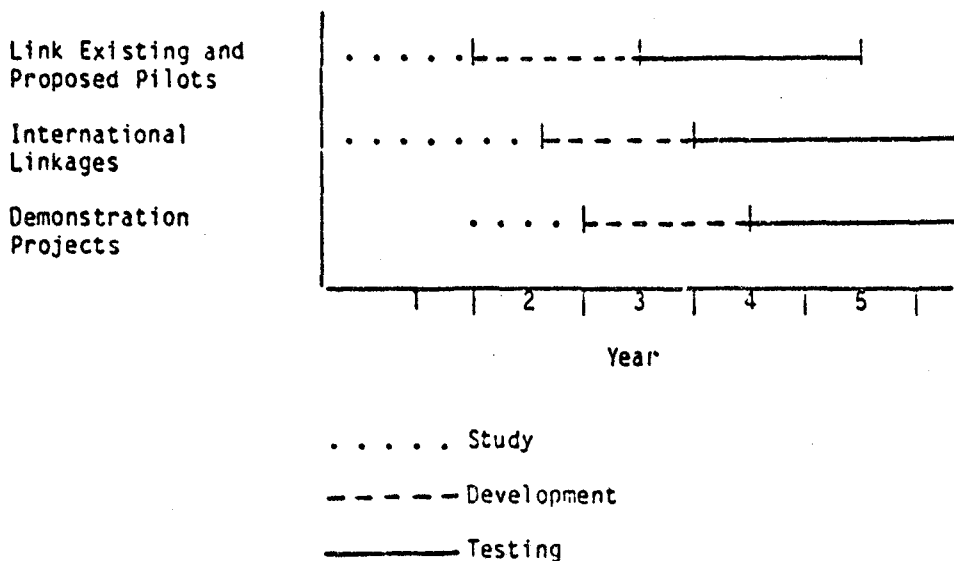appropriate.

Figure 6-1.   Proposed Phasing for a Global Resources Information
System Implementation

In conclusion, it should be emphasized that the material presented in
this report represents our interim thinking on the implementation of the GRIS
concept. Should Headquarters accept the threefold approach to GRIS implemen-
tation described above, we propose in the coming months to develop a more
detailed breakdown of the tasks required to support each program element.

The need for systems to facilitate the use of multisource data by multi-
disciplinary research terms to address global-scale scientific issues is great.
Systems must be developed to meet the needs of multidisciplinary research into
global problems. We feel the approach proposed herein addresses this need.

REFERENCES

1-1.   OSTA Commonality Analysis, Final Report, NASA-CR-169805, Vol. 1,
        OAO Corp., Beltsville, Maryland, June 1, 1981.

2-1.   "Space, A Resource for Earth," American Institute of Aeronautics
        and Astronautics (AIAA), 1290 Avenue of the Americas, New York
        10019, April 1977.

2-2.   "Land-Related Global Habitability Science Issues," NASA Office of
        Space Science and Applications, NASA Technical Memorandum 85841,
        June 1983.

3-1.   Data Management and Computation, Volume 1: Issues and
        Recommendations, Committee on Data Management and Computation,
        Space Science Board, National Research Council, National Academy
        Press, Washington D.C., 1982.

BIBLIOGRAPHY

1. *Applications Data Service Generic Requirements*, Computer Sciences Corporation, Report No. TM-79/6205.

2. *Applications Data Service 10-year Plan*, Second Draft for NASA review, Computer Sciences Corporation, CSC/TM-81/6106, September 1981.

3. des Jardins, R. (ed.), *OSTA Data Systems Planning Workshop Report*, Goddard Space Flight, May 1981.

4. des Jardins, R. (ed.), *Applications Data Service Study Report*, Goddard Space Flight Center, May 1981.

5. *Global Change: Impact on Habitability - A Scientific Basis for Assessment*, prepared by the Executive Committee of a workshop held at Woods Hole, Massachusetts, June 21-26, 1982. JPL Internal Publication D-95, Jet Propulsion Laboratory, Pasadena, California, July 7, 1982.

6. *OSTA Data Systems Planning Workshop Report, Part II: Technical Summary*, Draft 4, Applications Data Service Study Office, Goddard Space Flight Center, February 22, 1980.

N84
31748

UNCLAS

N84 31748

Dq

Appendix D

Three Papers on Application of Artifical Intelligence

and Remote Sensing

Tinney, L.R., C. Sailer, and J. E. Estes, 1983, "Applications of
Artificial Intelligence to Remote Sensing", *Proceedings of
Seventeenth International Symposium on Remote Sensing of
Environment*, Ann Arbor, Michigan.

Estes, J.E., C.T. Sailer, and L.R. Tinney, 1983, "User Analysis
Techniques", *Proceedings of National Telecommunications
Conference*, San Francisco, California.

Estes, J.E., C.T. Sailer, and L.R. Tinney, 1984, "The Potential
of AI techniques and Remote Sensing" *Professional
Geographer*.

1

THE POTENTIAL OF AI TECHNIQUES FOR REMOTE SENSING

John E. Estes*, Charlene Sailor* Larry R. Tinney

University of California
Santa Barbara, California

EG & G Energy Measurements Group
Las Vegas, Nevada

ABSTRACT

Remote sensing is a powerful tool for geographic analysis capable of producing large volumes of data in the spatial, spectral and temporal domains. A variety of "conventional" machine-assisted processing and analysis procedures have been developed in an attempt to improve the efficiency of information extraction from these data. Yet, in general the current state of the art of computer-assisted image analysis lags far 2 behind human interpretation techniques improved approaches to machine assisted image processing must be exposed. It is our conclusion that techniques adapted from the field of artificial intelligence (AI) can have significant, wide-ranging impacts upon computer-assisted remote sensing analysis.

AI-based techniques offer a powerful and fundamentally different approach to many remote sensing tasks.

In addition to computer assisted analysis AI techniques can also aid: "onboard" spacecraft data processing and analysis and database access and query.

Developing AI-assisted systems and analytical procedures will require a great deal of time and effort. The results,

1

particularly in the image processing area, should eventually

prove more accurate and robust than current statistical

approaches. In addition to the potential for increased

performance levels, AI-based research has focused upon user

acceptance issues and developed improved user interfaces that

should lead to greater acceptance of computer generated remote

sensing products.

INTRODUCTION

As Naisbitt (1984) states in this recent work Megatrends, we

have moved from an industrial to an information society.

Geographers today are faced with a wide variety of data as they

attempt to pursue their research. Figure 1 which focuses on

remotely sensed data as an information source illustrates the

variety of data types and sources and the need of users to tap

specific data of interest to their research. The information

explosion (depicted in figure 1) which has occurred in recent

years has lead to the increased use of automated procedures for

accessing, organizing, processing and analyzing spatial data.

Geographic information systems are responsive to the need for

improved capabilities to use spatial data for a wide range of

fundamental research and operational applications. It is

interesting to note that Naisbitt in Megatrends states that the

beginnings of this new information society dates back to 1956,

1957. (Naisbitt, 1984). This time frame corresponds to the

beginnings of satellite remote sensing with the launch of

Sputnik.

The Potential of AI

Remote sensing is a powerful data generation technology. A
technology which has in the past and will continue in the future
to have significant impacts on our ability to accomplish both
fundamental and applied geographic research (Estes, Jensen and
Simonett, 1980). Yet improved approaches to the processing and
analysis of remotely sensed data are required if we are to take
full advantage of the information potential inherent in the
products of this technology. In recent years a number of papers
and conferences have had as their theme the merging of remote
sensing and geographic information systems technologies (Shelton
and Estes, 1979; Shelton and Estes, 1981; and Estes 1981. This
merge is seen as a way to improve our ability to employ machine
assisted analysis procedures to improve our capability to extract
information from the vast array of complex spatial data available
to geographic researchers today. In these and more recent
articles, we have suggested a need To bring more "intelligence"
into the processing and analysis of data than is presently
permitted employing conventional statistical and mathematical
processing procedures. (Tinney, Sailer and Estes, 1983, Estes,
Sailer and Tinney 1983; Estes, 1984).

The current state-of-the-art of computer-assisted image
analysis generally lags far behind human interpretation
techniques. A comparison of these two methodologies illustrates
that the level of decision-making used in "automated" procedures
is still at a very low level, whereas analysts are capable of
taking advantage of higher level relationships between objects
both within a scene and across scenes.

For some applications existing computer-assisted procedures

produce classifications which are satisfactory for the task at hand. However, when faced with data sets which contain several different types of information, requiring substantial amounts of a priori knowledge and the application of reasoning or logic to solve the problem, conventional computer-assisted techniques are expremely limited. It is the opinion of the authors that much of the "technology transfer" problems associated with computer-assisted remote sensing using satellite-based systems are actually due to the inability of current techniques to systematically achieve accurate results.

Concepts from the field of artificial intelligence (AI) have proven useful in application domains with characteristics similar to those found in the field of remote sensing. From an examination of successful AI implementations, particularly in the area of "expert systems", it appears that there exists a potential for successful application to several remote sensing tasks. The area of most interest here, computer-assisted image analysis, is clearly one of the more difficult applications for which AI techniques are being applied. It appears evident, however, that AI-based techniques may be necessary if major improvements in automated image analysis performance are to be achieved. Current non-AI alternatives do not appear capable of ever reaching human analyst performance levels.

This paper provides a brief review of some basic AI concepts, including expert systems. Expert systems is an area of AI with the greatest potential for direct applications to various image analysis tasks. Introductory AI material is followed by

4

sections examining potential space and remote sensing
applications of AI technology. A major portion of this paper is
devoted to a comparison of human and computer-assisted image
analysis procedures. Following this comparison, some critical AI
implementation issues are addressed. The intent of this paper is
to provide the geographic community with an initial exposure to
AI techniques and convey a mix of excitement and caution
regarding their applicability to remote sensing tasks. We
believe that the impacts of AI techniques or the acquisition
processing and analysis of remotely sensed data can be profound,
but that their success implementation will require substantial
efforts.

ARTIFICIAL INTELLIGENCE

Artificial intelligence (AI) is the science of making
machines perform tasks that would require intelligence if done by
humans. A currently active subfield of AI is concerned with the
development of reasoning systems based upon formal logic systems,
such as the predicate calculus, for simulating human decision-
making. Decision models based on AI techniques are typically
expandable, easily altered, and extendable. AI employs computer
programs and programming techniques to elucidate the mechanics of
human thought processing in general, and human decision-making in
particular.

Artificial intelligence programs are fundamentally
different than conventional computer programs in a number of
ways. Table 1 lists some major differences between these types
of programs. Difference in programming strategy used is the most

basic distinction between AI and conventional programs. Non-AI
programs typically rely on the development of algorithms, which
are step-by-step methods for solving a particular class of
problems. A major problem with this approach is that there may
not be an algorithm that can solve every problem in a set of
problems for a particular application. In addition, even if an
algorithm exists, it may not be cost efficient for every problem
in that class of problems. AI programs model reasoning systems
using formal logics and employing heuristic searches. Heuristic
search can reduce the cost and circumvent the inefficiency of
exhaustive search techniques, and is of importance due to the
very large databases which the programs are typically called upon
to search. This could be particularly important in image
analysis where large spatial databases (e.g. geographic
information systems with multiple data planes), are often
required in the analysis task.

TABLE I. COMPARISON OF AI AND NON-AI PROGRAMS

AI PROGRAMS	versus	NON-AI PROGRAMS
o Heuristic		o Algorithmic
o Modular autonomous		o Knowledge
o Knowledge sources Flexible knowledge		o Inflexible stored knowledge deleted or augmented
o Symbolic		o Numeric
o °State' specific programming		o °Sequence' specific programming

In most AI programs there is a fairly rigid separation of input data concerning the problem; operations which manipulate the data; and, control structure which direct the operation. The knowledge base in an AI program is typically an autonomous component that can be directly modified, augmented or reduced by interactions with the operations portion of the program, or indirectly through the control structure. Knowledge is often stored in the form of rules, especially in "expert systems". In sharp contrast non-AI programs typically have rules contained within the control structure of the program, making the knowledge somewhat inflexible; such knowledge usually cannot be modified without physically rewriting the program.

Encoded knowledge of AI programs tends to be of a symbolic rather than numeric nature. Symbolic operations appear to more adequately characterize cognitive activities, such as problem-solving, planning and deduction (Duda and Shortliffe, 1983). AI languages have been developed with this data distinction in mind. An example of a language designed specifically for AI application is LISP, which manipulated wordlike objects (atoms) or groups of atoms (lists) rather than performing operations on a collection of data.

AI programs emphasize "state" specific programming. An action is performed when a specific set of conditions have been fulfilled, and not necessarily in any particular order. The control structure dictates what should happen in each 'state' but it is not necessary to know in advance when each state should occur. Non-AI programs dictate how the program will move from

one state into another, from one moment to the next. These
programs are considered to be sequence oriented. For additional
general information concerning artificial intelligence the
interested reader is referred to Shortliffe et al. (1979) and
Roberts (1971).

EXPERT SYSTEMS

Somewhat paradoxically, it has proven much easier to emulate
the problem-solving methods of some types of specialists than to
write more general programs that can approach a child's ability
to perceive, understand language, or make commonsense deductions.
Human experts are often distinguished by their possession of
extensive knowledge concerning a very specific domain of
problems. It is this very specificity of knowledge that has made
it feasible to develop "expert system" computer programs (Duda
and Shortliffe, 1983).

Expert systems are a subset of AI geared to problem solving
in a restricted domain of study. Expert systems take advantage
of the experiences, knowledge, and "rules of thumb" of experts in
a specified domain and call upon their expertise to solve
problems. Major components of an expert system are the knowledge
base (composed of modular knowledge sources) and the inference
mechanism which manipulates the knowledge contained in the
knowledge base. Performance of these systems is highly dependent
on the depth and structure of the knowledge base.

Classification programs tend to be the simplest and the most
successful expert systems (Duda and Shortliffe, 1983). Expert
consulting systems have been built that can diagnose diseases

The Potential of AI

consulting systems have been built that can diagnose diseases
(INTERNIST, Miller et al., 1982; MYCIN, Shortliffe, 1976),
evaluate potential ore deposits (PROSPECTOR, Duda et al., 1979)
suggest structures for complex organic chemicals (DENDRAL,
Lindsay, et al., 1980), and configure VAX computers (R1,
McDermott, 1982). Classification programs tend to be the
simplest and the most successful expert systems (Duda and
Shortliffe, 1983). Requirements for developing successful expert
systems include: the existance of at least one acknowledged
human expert that the expert's knowledge is based upon special
knowledge, judgment and experiencethat the expert(s) can explain
their approach to problem solving in their particular specialty
area; and, for successful implementation within the current
state-of-the-art, a narrow and well-defined area.

Human experts, however, do much more than just solve
problems. Their activities are typically characterized by a wide
range of behaviors with problem solving only being the most
evident. Experts are also able to explain results, learn and
restructure their knowledge, and, perhaps most importantly,
determine the relative importance of different elements within a
given situation. Flexibility of intelligence in an expert system
is largely derived from the range and levels of rules it
contains. A large number of rules is typically necessary because
the program can be faced with a wide range of situations. Expert
systems have already been developed which can approximate an
expert's performance in reaching a conclusion (Duda and
Shortliffe, 1983). However, much research is needed in the areas
of machine learning and knowledge restructuring.

The Potential of AI

Important operational characteristics found in many expert systems include the ability to: predict outcomes for events in their limited domain; utilize various types of data in reaching conclusions modify their knowledge base when confronted with conflicting assertions and "explain" their line of reasoning by providing the user with the pieces of knowledge used in reaching a conclusion. This ability to "explain" a line of reasoning has been important in user acceptance of these systems. In essence it adds what Naisbitt (1984) refers to as "high touch" to AI high tech approach. Another key characteristic affecting user acceptance of a computer system's results is the user interface. It is noteworthy that expert systems are almost always programmed with user interactions in mind (geologists speak as geologists, chemists as chemists, etc.). Indeed, "natural language" processing is an important area of AI research and this greatly enhances the likelihood of improved user interfaces in future systems.

Additional introductory material concerning the development of expert systems can be found in Duda and Gaschnig (1981) and Stefik et al. (1982).

REMOTE SENSING APPLICATIONS OF ARTIFICIAL INTELLIGENCE

One could probably find as many potential applications of AI techniques to the field of remote sensing. The focus of this discussion will address earth resource survey applications of remote sensing. Furthermore, our emphasis is upon image oriented data analyses.

The Potential of AI

Among generic AI research areas of special interest are:

o   vision, perception, and pattern recognition

o   problem-solving; and,

o   information storage and retrieval.

Less central, but also of interest, are automated speech recognition and synthesis.   These topics are being investigated as a means to allow voice data entry by image analysts (Lukes, 1983).

Some remote sensing applications appear especially well suited for potential use of AI techniques. We will discuss three broad and somewhat overlapping application areas. Our discussion is of the use of AI in these applications areas is necessarily somewhat speculative by necessity, given the present lack of examples. Application areas discussed are: intelligent "onboard" processing advanced data base query and interrogation, and automated image analysis.

INTELLIGENT ONBOARD PROCESSING AND ANALYSIS

Even current remote sensing platforms must often accomplish a great deal of onbaoard processing 'e.g. recording and transmission of imagery to ground stations).   When the platforms are unmanned satellites this processing usually must be accomplished in a semi-autonomous manner. This is especially true for platforms containing sensor systems which are travelling in deep space. But even in earth orbit significant benefits could be obtained from more intelligent onboard processing.

Examples of advanced onboard processing include: automated

11

navigation and locating of scenes to be imaged; screening of
conditions that might influence data collection, and; automated
change detection. Data compression could also fall into this
category if adaptive procedures are used to select optimal
compression parameters or techniques.

An example of research oriented towards intelligent onboard
processing is the Feature Identification and Location Experiment
(FILE), which flew on the second Space Shuttle mission Silvertson
et al., 1982). The FILE is a system designed to test techniques
for real-time autonomous classification of four primary earth
features: water; vegetation; bare land; and clouds, snow, and
ice, clouds snow and ice are considered one category on file as
all of these features typically exhibit high reflectance.

The FILE system senses earth radiation in two spectral bands
(red and infra-red bands centered at 0.65 and 0.85 microns,
respectively). Real-time classification decisions are made by
the instrument based on a predetermined partitioning of the two-
band measurement space. Viewing angle factors (solar zenith
angle, solar sensor azimuthal angle, and sensor viewing angle)
are accounted for by a sunrise detector subsystem and specific
viewing angle assumptions. Variations in atmospheric conditions
are presently not directly accounted for; however, a ratio
technique is used that to some degree minimizes atmospheric
effects.

An obvious application for a FILE-type system would be for
onboard data acquisition decisions to be based upon cloud cover
conditions. General or location specific thresholds could be

used as a basis for the system activation of an imaging sensor
system. A more sophisticated version could evaluate atmospheric
haze conditions based on a combination of information from
additional sensors and through the interrogation of ground
stations. criteria.

Given information of approximate platform location and
orientation parameters, in conjunction with a reference image of
the area involved, an advanced systems could perform automated
image cross-correlation. The calculated offsets could then be
used for precise navigation and scene selection purposes (perhaps
using pointable sensors) or as a basis for an onboard change
detection system (using image differencing techniques). For
change detection purposes further processing would, however, be
required to separate changes due to natural phenologic cycles
from those caused by unusual episodic events, but the potential
benefits of an intelligent data acquisition system such as this
appears substantial (Freitas and Gilbreath [eds.]. 1982).

DATABASE QUERY AND INTERROGATION

Management of imagery and other types of spatial data can be
critical to remote sensing projects. As spatial data sets become
larger the tasks involved in processing and analysis become more
difficult and often more important. Among the key tasks of
remote sensing database management are: identifying existing
remotely sensed data germaine to the task; identifying existing
collateral data; selecting what new data is required; and,
providing user friendly interaction. When remote sensing and
other types of spatial data (e.g. maps) are integrated for

analysis it is commonly done within the framework of a geographic
information system (GIS). It is typical for a GIS to offer some
level of database management capabilities. Few existing GIS,
however, have database management facilities comparable to those
of well established commerical database systems (Blaser [ed.]
1980, Klinger et al. [eds.], 1977). Commercial applications of
database management systems (DBMS) have been extremely successful
in banking and other accounting oriented applications. It is in
these types of DBMS that most of the advanced work has been
accomplished. However, these commercially-oriented are not
easily modified to serve in spatial data manipulation. As a
result, the development of advance spatial data base management
systems has suffered and indeed many existing GIS packages do not
even exploit the current level of DBMS sophistication.

For some time, database systems technology has been an
active area within computer science. Intelligent retrieval from
databases has also become one of the major research topics within
the artificial intelligence community (Nilsson, 1980). Some of
the most advanced DBMS offerings now incorporate AI techniques in
their user interfaces, permitting limited natural language
intereactions (e.g. some can respond to limited vocabulary in
spoken English).

It has recently been suggested that GIS technology may stand
to significantly gain should certain methods from database
systems and artificial intelligence be incorporated (Estes,
1982). An excellent review of these topics is presented by
Pazner et al, (1983). Stressing the urgent need to improve GIS

technology, Pazner notes that it is spatial datasets in general,
and the very large georeferenced databases in particular, which
stand to benefit most from the incorporation of Artificial
Intelligence techniques. Pazner goes on to say that "Due to the
considerable complexity of handling very large databases, and
spatial ones in particular, it appears inevitable that AI methods
shall be integrated into GIS systems in the future."

There are three obvious ways by which AI-based DBMS
methodologies may improve future GIS performance:

o   approved user interfaces to facilitate user
     interactions with the system

o   increased efficiency for querying and altering data
     according to the user's needs

o   general enhancement of capabilities to allow
     the system to answer a wider repertoire of
     questions about the data as well as about itself

The interested reader is referred to the review by Pazner et
al. (1983) for a more detailed discussion. As an outgrowth of
this work researchers at the University of California are
currently developing a prototype GIS which incorporates AI
techniques. The system being developed under the direction of
Drs. Terence Smith, Donna Peuquet and John Estes has an advance
structure with parallel object and attribute trees. The system
has the capability to learn from queries put to it by users.

AUTOMATED IMAGE ANALYSIS

This section examines the potential application of
artificial intelligence techniques to automated image analysis
tasks from the perspective of earth resources remote sensing

applications. As background we first briefly compare and contrast the current status of human and computer-assisied approaches to image interpretation based on research by Estes et al, 1983 and Tinney, 1983. A brief review of automated image analysis procedures is then presented to demonstrate the recent trend toward procedures which more closely mimic those of human analysts.

The tasks of both a human image interpreter and automated image inter-pretation are basically similar, namely detection, identification, measurement, and problems solving (see Figure 2). In most image interpretation problems of interest to geographers computer-assisted techniques currently lag far behind human techniques in terms in both speed and accuracy. Continual improvements in computing hardware, especially lower cost array processor well suited for image processing, are rapidly improving the speed of computer-assisted image analysis.

Substantial improvements in accuracy, however, may not be possible without fundamental changes in the structure of automated procedures.

One of the most significant contributions to date of the computer-assisted approach has been some necessary focus upon the basic elements and processes involved in image interpretation. Because all image operations must be explicitly specified for a computer program it has been necessary to move away from the ill-defined "gray box" mode of operation all too often prevalent in discussions of human image interpretation processes.

Early automated work emphasized pattern recognition approaches. A substantial subfield has developed within computer

16

science aimed at developing the ability of computers to
"recognize" patterns of data. When these techniques were first
invented, the mathematical elegance of the basic ideas was so
attractive that many early researchers were overly ambitious and
optimistic about the general utility of the techniques. It is
now recognized that pattern recognition methods alone are
inadequate in situations which require awareness of context or
the use of knowledge, characteristics common to most interesting
image analysis problems. The AI field has addressed these types
of problems and the techniques already developed promise to
substantially change our automated procedures; in many instances
these changes will result in techniques similar to those employed
by human image analysts, thus acceptance of these new techniques
should be rapid once they are more fully developed.

IMAGE ELEMENTS

The primary elements of image interpretation (tone,texture,
site,etc.) appear common to both human and computer-assisted
procedures. (Figure 2) Their current usage, however, is
dramatically different. Each of the elements can readily be used
by image analysts in performing their tasks. In contrast, most
automated remote sensing image analysis is based on tone
or color exclusively (here we are considering multispectral
response as an extension of color) (see figure 5). This seems to
be the result of an overly focused pursuit of the "multi-spectral
signature" concept and the relative simplicity of implementing
"single pixel" multispectral classification algorithms using

17

statistical pattern recognition techniques. At some point, and
hopefully soon, the remote sensing community, needs to recognize
the limits of the multispectral concept and pursue a broader
based approach to scene analysis.

Various procedures have been developed to incorporate
additional elements but only texture and site currently appear to
be used on an operational basis. As used in automated analyses,
however, textural features tend to be simplistic representations
that certainly do not capture all of the spatial information that
human image analysts refer to as "texture".

Factors relation to site typically are used via
stratification techniques and/or class a priori probabilities.

Automated use of other image interpretation elements appears
to have been demonstrated, if only on a limited basis. Related
to site as a locational element is associationg/context. Several
recent examples of "contextual" classification make explicit use
of neighboring pixel characteristics, thus incorporating some
level of association information. Various "region growing"
algorithms have been developed that could allow region sizes to
be determined and used as a discriminating variable. Shape is a
very difficult element to incorporate but methods have been
defined to use shape albeit on a limited basis to date (e.g. via
a syntactical classifier, to be discussed later). Cloud and
shadow relationships have been used as a basis for detecting
clouds; the use of additional sun angle information could be used
to determine cloud height. It should be stressed that even
though use of most elements has been demonstrated, many examples
entail only a small aspect of each element's total information

18

content. And the majority of current computer-assisted image
analysis still relies solely upon single pixel multispectral
response (tone/color).

Edge detection and segmentation are two procedures growing
in use in the computer-assisted approach to image interpretation.
The necessity of using these procedures to define some elements
(e.g. size and shape) identifies a possible gap or illogical
structure in our set of basic elements as currently defined for
human image interpretation. It seems likely that a hierarchy of
elements and operators exists. Even more disturbing is that some
of the texts on image interpretation present different sets of
basic interpretation elements; imagine chemistry with a shifting
set of basic chemical elements! It is clear that the basic
elements of image interpretation need to be carefully reviewed
and defined.

TRAINING AND LABELING PROCEDURES

Scene specific labeling and training by example is widely
used in multispectral pattern recognition. Little "signature
extension" capability, (the effective use of training or labeling
data from one scene for another scene), has been demonstrated.
This is largely due to environmental variations (e.g.,
atmospheric conditions, climate zones, etc.). Training by
example is directly analogous to the use of inter-pretation
"keys" by image analysts. Current digital pattern recognition
techniques, however, are not as flexible as those employed by an
analysts, who can mentally account for slight temporal offsets,

atmospheric variations, and other factors that are beyond the capability of most computerized algorithms.

An alternative to scene specific training or labeling is the development of a mathematical model protraying the object or phenomenon of interest. Recent work by Badwar (1981) clearly demonstrates this approach using multidate sets of Landsat imagery for crop identification purposes. Badwar has developed crop specific mathematical models of "greeness" (a transformation of Landsat data indicative of plant biomass) through time. Curve matching techniques are subsequently used which allow for some temporal offsets (signature extenstion in the temporal domain). This approach appears to perform well for a variety of environments (spatial signature extension). The mathematical model approach appears somewhat analogous to the perceptual models employed by human analysts, although no deductive reasoning is explicitly incorporated.

INCORPORATING COLLATERAL DATA

Collateral material may be used for computer-assisted image analysis in basically the same ways used by image analysts, although the methods for incorporating such data are less well developed. Hutchinson (1982) and others have systematically ... .ered the various methods whereby both discrete (e.g. soils maps) and continuous (e.g. terrain slope) collateral data may be incorporated into pattern recognition algorithms. Implicit in some of these methods is the use of collateral data in a geographic information system framework that allows simple logic

the development of analysis procedures that extend beyond simple pattern recognition and into the realm of symbolic or logical reasoning techniques which are characteristic of Artificial Intelligence.

CLASSIFICATION TECHNIQUES

Available classification techniques has evolved substantially since the first digital classification algorithms were developed. However, many of the original algorithms, are still widely used due to the computational complexity of more advanced techniques. The following section compares various approaches to multispectral classification focusing on their strengths, limitations and future applicability to advances in computer-assisted image analysis. The classificiation approaches examined are statistical pattern recognition, syntactical pattern recognition, and symbolic reasoning.

Statistical Pattern Recognition

The most common computer-assisted classification procedure is also among the most primitive; what is often termed "statistical pattern recognition" makes use of training data to characterize patterns of interest in some statistical manner. This decision-making method has been found to be most effective in two types of problems (Raphael, 1976):

1. Classification of complex signals when the proper features are measured and the number of dimensions are kept small (typically less than ten).

2. Recognition of simple shapes.

21

The Potential of AI

One result of this is a major effort towards optimal
representation and feature extraction (e.g. ,band selection and
various transformations, such as a principal components, directed
towards dimensionality reduction). As noted earlier, however,
pattern classification methods alone are virtually useless in
situations which require awareness of context or the use of
pertinent additional knowledge.

The use of decision tree structures is sometimes used to
improve classification results. This is often termed the
"layered" approach to classification. This approach clearly
implies the use of higher order, heirarchical decision making
procedures. The integrated use of image analysis in a geographic
information systems context has provided a major impetus to this
approach (Hallada et al., 1981). In many instances it is
possible to implicitly incorporate high order inference rules
into a tree structure. That this approach typically requires the
sometimes tedious construction of a new tree structure with each
new problem or data set has led to movement towards procedures
that make more explicit use of inference rules.

Syntactical Pattern Recognition

From the study of languages comes another slightly higher
order procedure termed syntactic pattern recognition.
Conventionally, a language is defined as a set of strings over an
alphabet, where the alphabet consists of the set of all symbols
which can appear in the language strings. A string is a definite
ordered sequence of symbols. A grammer is a set of rules which
define how the strings of the language are formed. A grammar can

be used to recognize the language's strings by using the rules in
reverse order. This concept can be generalized in a number of
ways to define grammars for classes of images. This approach has
been most widely used in image analysis to recognize shapes based
upon order of component parts. Syntactic methods have been used
for locating highways and rivers in Landsat images and for
texture modeling.

SYMBOLIC REASONING

It seems appropriate here to emphasize the necessity of
exploring the development of symbolic reasoning procedures that
employ formal inference. It is interesting to note that this
approach to image analysis is being vigorously pursued by
computer scientists interested in computational approaches to
"image understanding" (3rady, 1982). Much of this work has been
conducted under the Defense Advanced Research Project Agency's
Image Understanding Program. Although most of this research is
directed towards high resolution panchromatic imagery, the
techniques are analogous to those employed by human analysts and
will probably need to be pursued if substantial progress is to be
made at extracting the inherent information of higher order image
elements. Although these techniques are not presently
incorporated within the field of earth resource remote sensing it
appears likely that symbolic reasoning approaches, or some
derivative thereof, will eventually dominate computer-assisted
image interpretation.

In 1975, the Defense Advanced Research Project Agency
(DARPA) initiated a major research effort called the Image

Understanding Program (Druffel, 1979). This program was the
direct outgrowth of an earlier effort called the image Processing
Program through which research in image enhancement, image
restoration, encoding for bandwidth compression, and visual
system modeling was sponsored. A major purpose of the image
Understanding Program was to investigate the use of a priori
knowledge to facilitate an understanding of the relationships
among objects in a scene. Primary thrusts included smart
sensors, iconics (visual phonetics) and symbolic representation.
Potential applications of image understanding research are
discussed in Druffel (1979), Duda and Garvey (1980), and Lukes
(1981). An excellent overview of recent theoretical developments
in the program is provided by Brady (1982). Brady identifies
some common themes that appear to have crystalized over the past
decade as follows: representations have been developed that make
explicit the information computed by a module (most modules do
not work on raw data, but on computed representations, e.g. edge
images); the mathematics of image understanding are becoming more
sophisticated; locally parallel architectures have been
developed; attention has shifted from restrictions on the domain
of application of a vision system to restrictions on visual
abilities, with current concentration topics corresponding to
identifiable modules in the human visual system; and, there are
growing links between image understanding and theories of human
vision.

Image representations are discussed in some detail by Ahuja
and Schacter (1982), who stress the importance of the concept

The

Potential of AI

using the framework of an "image model". They distinguish
between low-level models that provide concise abstractions of
spatial variations (e.g. auto-correlation function, variogram,
mean, gradient, etc.) and high level models that invo: ·e highly
semantic descriptions and a large degree of outside of a priori
knowledge about a scene to augment the raw data.

Of the many disciplines closely related to image
understanding, four appear of particular interest to the computer
science community: image processing (image transmission,
storage, enhancement, restoration), computer graphics (display of
visual information), computer-aided design and manufacture or
CAD-CAM (requires attention to surface representation), and
pattern recognition. Pattern recognition historically is the
area most closely related to image understanding. Brady (1982)
identifies the most significant differences between pattern
recognintion and image understanding as follows: pattern
recognition systems are typically concerned with recognizing the
input as one of a small set of possibilities whereas image
understanding aims to construct rich descriptions that cannot
be enumerated in advance but need to be constructed for each
individual image; pattern recognition systems are mostly
concerned with two-dimensional images; image understanding has
dealt extensively with three-dimensional images; and, most
significantly, pattern recognition systems usually operate
directly on the image while image understanding and indeed most
visual processes operate not on the image but on symbolic
representations that have been computed by earlier processing
(e.g. edge detection).

25

The Potential of AI

It seems apparent from this review that AI-based approaches offer a logical means to pursue more advanced image analysis. Not so clear are the specific tasks and limitations associated with this approach. The following section addresses some of these issues.

DEVELOPING AI ASSISTED IMAGE ANALYSIS SYSTEMS

Developing an AI-based image analysis system is a complex task typically requiring years of effort. Although development periods have been speeded up with the advent of new "knowledge acquisition" tools, the development of a successful expert system is still considered to fall in the two year-plus range (Davis, 1982). Image analysis is one of the more difficult areas for which AI techniques are being applied.

The field of AI is still relatively small, thus there is going to be competition for the limited amount of knowledge-engineering expertise. Recent successes in AI are still more promising than productive, but the promise for improvement has brought about a great deal of interest by disciplines that are comparatively "rich" when compared to earth resources remote sensing (medicine, defense, petroleum, etc.).

The computational resources considered necessary for developing AI systems is also different than those typically available for remote sensing. Specialized languages such as LISP and PROLOG are strongly preferred by most AI researchers. Even specialized hardware (e.g. "LISP machines") is becoming available to enhance the performance and development of AI programs.

26

The Potential of AI

There is also the need to pursue only appropriate tasks
tasks to codify the knowledge base necessary for these tasks in
the systems. This will require substantial work and obviously a
critical review of how image analysts go about doing their job,
for if we cannot formalize how analysts go about their tasks we
cannot automate the procedure.

Although difficult, the process of automating image analysis
using AI techniques has already begun for applications outside
the realm of earth resources. Among the initial systems in this
area are the ACRONYM system at Stanford University (Binford and
Brooks, 1979), the HAWKEYE system of Stanford Research Institute
(Bolles et al.,), and the ARGOS system of Cargegie-Mellon
University (Rubin, 1978). A very good book discussing the
development of such a system has been written by Nagao and
Matsuyama (1981).

Finally, special attention will need to be directed towards
those conditions that are perhaps somewhat unique to geography in
particular and earth resources applications in general. For
example, the uses of AI applications where a multistage remote
sensing approach combines data from literature and archives with
data from the field, and imagery from sensor systems cover larger
and larger areas in a geographic information system framework.
Indeed as seen above work which combines image analysis and
information systems assisted by AI techniques is a major key to
the future of large scale remote sensing applications. These
potential implementation issues suggest that the development of
AI-based systems within the remote sensing community will not
occur as rapidly as previous developments. The tasks ahead are

27

both difficult and time consuming. Careful attention will be
required to assure the "right" tasks are automated and a proper
balance of man-and machine-oriented procedures is established.

CONCLUSIONS

Many of the techniques being developed by the artificial
intelligence community appear relative to the needs of remote
sensing. From improved classification performance to better user
interfaces, the potentials offered by AI-based techniques are
promising and could aid significantly in facilitating more
widespread acceptance of computer generated remote sensing
products.

As reviewed in this paper, the basic foundations of human
image interpretation as currently practiced in earth resource
applications of remote sensing are in need of critical review;
the basic elements are ill-defined and perhaps incomplete. The
development of better computer-assisted procedures for image
analysis will require substantial efforts even after we better
understand our own procedures. AI-based procedures seem to be
appropriate for improving image analysis but conditions somewhat
unique to our field will need to be carefully addressed.

As seen in figure 4, as aerial photographic interpretation
has evolved into the discipline of remote sensing, the field has
become increasingly complex. Sensor platforms have moved from
tethered balloons to orbiting satellites. Sensor systems have
developed from cameras which employ analog processes to produce
photographs to electronic systems which digitally record images
of a given scene. Analysis techniques have progressed from the

28

use of manual image analysis procedures to perform simple
identifications of objects or phenomena and their significance
from photography alone of a small area; to the application of
machine-assisted analysis in the modelling of complex processes
employing multiple data sources for global surveys.

The reader has only to examine the current edition of The
Manual of Remote Sensing to get a feel for this complexity
(Colwell, 1983). This two volume work of 2440 pages by over 200
authors details the ranges of systems techniques and applications
of remote sensing today. Yet, many of use in remote sensing and
many who want to apply remote sensing to their problems are
frustrated by this complexity of: what systems to choose from;
what times to collect data or where to get data; how to process
it or; what has been done to it before the researcher gets it.
These are but a few of the valid questions researchers or
applications oriented users may pose. There is a need for ways
of helping the users of remote sensing data through this process.
What is needed is a way to unlock the complex "black box" of
remote sensing and make this high tech field more accessible to
the users such as regional, historical, cultural and physical
geographers who can make excellent use of these data in their
given research areas. As Naisbitt (1984) would say, high touch
is required. High touch must accompany high technology to
counter balance the normal human response to forced technology.
Often in the 1960's and 1970's potential users of remote sensing
were "turned off" by oversell of this technology. This was
forced technology. Indeed this situation is still occurring.

Yet, in a subtle way, the increased application of techniques from the field of AI is a trend towards adding the high touch required to increase the potential acceptance of remote sensing by a wider variety of users. The work in expert systems and natural language is particularly important here. This work is sharing the process of acquisition, processing and analysis of remotely sensed data easier and more understandable to users at a variety of levels of sophistication. This AI high touch can, in effect, if properly implemented, let geographers function more as geographers and less as librarians, communications specialists, computer scientists, image processors and so on. AI may just hold an extremely significant key to unlocking the potential of remote sensing for not only the broader geographic community specially, but for scientists and users in many other fields in universities, government agencies and private industry as well. Research on the applications of AI in remote sensing must continue and be expanded.

REFERENCES

Ahuja, N., B.J. Schacter, 1981, "Image Models", Computer Survey, Vol. 13, No. 4, pp. 373-397.

Badwar, G.D. and K.E. Henderson, 1981, "Estimating Development Stages of Corn From Spectral Data--An Initial Model", Agronomy Journal, Vol. 73, pps. 748-755.

Blaser, A., [9], Blaser, A., (Ed.), 1980, "Database Techniques for Pictorial Applications" in Lecture Notes in Computer Science, Vol. 81.

Binformd, T.O. and R.A. Brooks, "Geometric Reasoning in ACRONYM", 1979, Proceedings of a Workshop on Image Understanding, pp. 48-54.

Bolles, R.C., L.H. Quam, M.A. Fischler and H.C. Wolf, 1978, "The SRI Road Expert: Image-to-Database Correspondence", Proceedings of a Workshop on Image Understanding, pp. 163/174.

Colwell, R.N. (ed. in chief) 1983 Manual of Remote Sensing 2nd edition, Falls Church, Va. American Society of Photometry 2440 p.

Davis, R., 1982, "Expert Systems: Where Are We? And Where Do We Go From Here?", The AI Magazine, Spring, pp. 3-22.

Druffel, L.E., 1979, "Derivation of Information from Images", Proceedings of the Society of Photo-Optical Instrumentation Engineers, Vol. 205, Image Understanding Systems II, pp. 53-57.

Duda, R.O. and T.D. Garvey, 1980, "A Study of Knowledge-Based Systems for Photo Interpretation", SRI International, U.S. Army Engineer Topographic Laboratories Report ETL-0235.

Duda, R.O. and J.G. Gaschnig, 1981, "A Study of Knowledge-Based Systems for Photo Interpretation", SRI International, U.S. Army Engineer Topographic Laboratories Report ETL-0235.

Duda, R.O. and J.G. Gaschnig, 1981, "Knowledge-Based Expert Systems Come of Age", Byte Magazine, pp. 238-281.

Duda, R.O. and E.H. Shortliffe, 1983, "Expert Systems Research", Science Volume 220, Number 4594, pp. 261-268.

Estes, J.E., 1981, "Remote Sensing and Geographic Information Systems Coming of Age in the Eighties", Proc. of the Seventh Pecora Symposium.

Estes, J.E., 1984. "Improved Information Systems: A Critical Need", Proc. Tenth Machine Data Processing Symposium. West Layfette, Ind. Purdue University Laboratory for Applications of Remote Sensing (in press).

Estes, J.E., J.R. Jensen and D.S. Simonett, 1960. "Impact of Remote Sensing on U.S. Geography", Remote Sensing of Environment, Vol. 10, No. 1, pp. 43-80.

Estes, J.E., C.T. Sailer and L.R. Tinney, 1983. "User Analysis Techniques", Proceedings of the National Telecommunications Conference, San Francisco, California.

Freitas, R.A. and W.P. Gilbreath. (Eds.), 1982, "Advanced Automation for Space Missions", NASA Conference Publication 2255, Proceedings of the 1980 NASA/ASEE Summer Study.

Hallada, W.A., F.C. Mertz, L.R. Tinney, M.J. Cosentino, J.E. Estes, 1981, "Flexible Processing of Remote Sensing Data Through Integration of Image Processing and Geobased Information Systems", Proceedings of the Fifteenth International Symposium of Environment, ERIM.

Hutchinson, C.F., 1962, "Techniques for Combining Landsat and Ancillary Data for Digital Classification Improvement",

pp. 123-130.

Klinger, A., K.S. Fu, T.L. Kunii (Eds.), 1977, Data Structures, Computer Graphics, and Pattern Recognition, Academic Press.

Lindsay, R.L., B.G. Buchanan, E.A. Feigenbaum, and J. Lederberg, 1980, Applications of Artificial Intelligence for Organic Chemistry: The DENDRAL Project, McGraw-Hill.

Lukes, G.E., 1981, "Computer-Assisted Photointerpretation Research at United States Army Engineer Topographic Laboratories (USAETL)", Proceedings of SPIE, Volume 281-Techniques and Applications of Image Understanding, pp. 85-94.

Lukes, G.E., in press, "CAPIR Computer-Assisted Photo Interpretation Research", Proceedings of the Seventeenth International Symposium on Remote Sensing of Environment. ERIM.

MCDermott, J., 1981, "R1 The Formative Years", the AI Magazine, Summer pp. 21-29.

Nagao, M. and T. Matsuyama, 1981, A Structural Analysis of Complex Aerial Photographs, Plenum Press.

Miller, R.H., M.L. Minsky, and D. Smith, 1982, "Space Applications of Automation, Robotics and Machine Intelligence Systems ("ARAMIS"), MIT Space Systems Laboratory and Artificial Intelligence Laboratory, Phase 1, Final Report.

Miller, R., H. Pople, Jr., and J. Myers, 1982, "INTERNIST-1", New England Journal of Medicine, Volume 307.

Nagao, M. and T. Matsuyama, 1981, A Structural Analysis of Complex Aerial Photographs, Plenum Press.

Nilsson, N.J., 1980, Principles of Artificial Intelligence, Tioga Press.

Pazner, M., T. Smith, and J. Estes, 1983, "Improved Querying of Large Spatial Databases Utilizing Artificial Intelligence Techniques", University of California at Santa Barbara, Remote Sensing Unit, USDI Contract #14-08-0001-21162.

Roberts, S.K., 1981, "Artificial Intelligence", Byte Magazine, September, pp. 164-178.

Rubin, S.M., 1978, "The ARGOS Image Understanding System", Proceedings of a Workshop on Image Understanding, pp. 159-162.

Shelton,R.L. and J.E. Estes, 1979. "Integration of Remote Sensing and Geographic Information Systems", Proc. of the 13th International Symposium on Remote Sensing of the Environment, Environmental Research Institute.

Shelton, R.L. and J.E. Estes, 1981. "Remote Sensing and

Geographic Information Systems: An Unrealized Potential", Geo-Processing, Vol 1, Elsevier Scientific Publishing Co., Amsterdam, pp.395-420.

Shortliffe, E. H., "Computer-Based Medical Consultations: "MYSIN", Elsevier/North-Holland.

Shortliffe, E.H., B.G. Buchanan, and E.A. Feigenbaum, 1979, "Knowledge Engineering for Medical °Decision Making: A Review of Computer-Based Clinical Decision Aids", Stanford University, Computer Science Department, STAN-CS-79-723.

Silvertson, W.E. Jr., R.G. Wilson, G.F. Bullock, R.T. Schappell, 1982, "Feature Identification and Location Experiment", Science, December, Vol. 218, pps. 1031-1033.

Stefik, M., et al., 1982, "The Organization of Expert Systems, A Tutorial", Artificial Intelligence, Vol. 18, pps. 135-173.

Tinney, L.R., 1982, "Toward an Improved Understanding of the Image Interpretation Process", Masters Thesis, University of California, Santa Barbara.

Tinney,L.R., C.T. Sailer and J.E. Estes, 1983. "Applications of Artificial Intelligence to Remote Sensing", Proceedings of the 17th International Symposium on Remote Sensing of Environmunt, Ann Arbor, Environmental Research Institute of Michigan.