Key Issues in the Analysis of Remote Sensing Data

A Report on the Workshop
June 22-23, 1981
Purdue University
West Lafayette, Indiana

by Philip H. Swain

LARS Technical Report 062481

Purdue University
Laboratory for Applications of Remote Sensing
West Lafayette, Indiana 47907 USA

KEY ISSUES IN THE ANALYSIS OF REMOTE SENSING DATA

A Report on the Workshop

June 22-23, 1981

Purdue University

West Lafayette, Indiana

By

Philip H. Swain

LARS Technical Report 062481

Purdue University
Laboratory for Applications of Remote Sensing
West Lafayette, IN 47907

Star Information Form

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Key Issues in the Analysis of Remote Sensing Data		5. Report Date June 1981
		6. Performing Organization Code
7. Author(s) Philip H. Swain		8. Performing Organization Report No 062481
9. Performing Organization Name and Address Purdue University Laboratory for Applications of Remote Sensing 1220 Potter Drive West Lafayette, IN 47906-1399		10. Work Unit No.
		11. Contract or Grant No. NAS9-15466
12. Sponsoring Agency Name and Address NASA Johnson Space Center Remote Sensing Research Division Houston, TX 77058		13. Type of Report and Period Covered Technical
		14. Sponsoring Agency Code
15. Supplementary Notes		

16. Abstract

A workshop was convened at Purdue University on June 22-23, 1981, for the purpose of assessing the state-of-the art of machine analysis of remote sensing data and determining the key research problems remaining as barriers to broader and more effective use of this technology. This report is the proceedings of that workshop.

17. Key Words (Suggested by Author(s))	18. Distribution Statement		
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages	22. Price

*For sale by the National Technical Information Service, Springfield, Virginia 22161

NASA - JSC

KEY ISSUES IN THE ANALYSIS OF REMOTE SENSING DATA

A Report on the Workshop
June 22-23, 1981
Purdue University
West Lafayette, Indiana

Philip H. Swain

INTRODUCTION

The year 1981 found the remote sensing community assessing the results of completed applications-oriented tests of the remote sensing technology and looking ahead with great anticipation to new opportunities for advancing the technology and broadening its use. For example, the Large Area Crop Inventory Experiment (LACIE)[1] and the development of a Forest Resources Information System (FRIS) for commercial application[2] had demonstrated the capabilities and limitations of the mid-1970's technology. The future availability of new sensors, including the Thematic Mapper and the French SPOT multispectral sensor, plus the anticipation of renewed research support from NASA through a new fundamental research program provided motivation for understanding clearly both the current status of the technology and the directions which future research must take to best utilize remote sensing.

These considerations stimulated the convening of a Workshop on Key Issues in the Analysis of Remote Sensing Data at Purdue University, June 22-23, 1981, in conjunction with the 1981 Symposium on Machine Processing of Remotely Sensed Data. Jointly sponsored by Purdue's Laboratory for Applications of Remote Sensing (LARS) and NASA, the workshop had the following objectives:

1. To assemble experts in remote sensing and related information-processing and image-processing technologies for the purpose of making an up-to-date assessment of the state-of-the-art of machine analysis of remote sensing data.

2. To determine the nature of the key research problems remaining as barriers to broader and more effective use of machine analysis of remote sensing data.

3. To produce a report for use by interested researchers and potential research sponsors detailing the findings and recommendations of the workshop participants.

To achieve these objectives, invitations to participate in the workshop were extended to several well-established scientists and engineers in the field from universities, research institutions, and government. The workshop also was publicized in the widely mailed preliminary program of the Machine Processing Symposium. Thirty-six participants were on hand when the workshop was called to order. (See Appendix 1.)

To establish a common point of departure for the meeting, the report entitled "Basic Research Planning in Mathematical Pattern Recognition and Image Analysis," by Jack Bryant and L.F. Guseman of Texas A&M University[3], was mailed to those who registered in advance and was distributed at the conference to all others who registered. The report summarized the conclusions of a NASA-commissioned working group charged with defining a fundamental research program in image processing for remote sensing. As such, it provided a natural starting point for the discussions planned for the workshop.

Sessions of the workshop (see Appendix 2 for Workshop Schedule) focused on:

* Data Bases and Image Registration, including presentations on Data Bases for Remote Sensing, Image Preprocessing Operations, and Map-Oriented Considerations.

* Advanced Technology, including presentations on Advanced Digital Systems, and Artificial Intelligence Methods.

* Information Extraction, including presentations on Classification, and Classifier Training Considerations.

Each session had a reporter assigned to record and summarize key points in the presentations and the associated discussion periods. (The proceedings compiled by the session reporters may be found in Appendix 3 of this report.) The workshop ended with general comments from Mr. R.B. MacDonald of NASA/JSC, representing the workshop cosponsor, concerning the near and intermediate term outlook for support of fundamental research in remote sensing.

SUMMARY OF WORKSHOP PROCEEDINGS

With regard to data bases and image registration, it was surprising to find a great deal of disagreement on the degree to which improved registration and rectification of data are required. There seemed to be a general consensus that research is needed in

* Improved platform control and sensor modeling to reduce the need for rectification and registration.

* Modeling atmospheric effects and the atmosphere point spread function.

* Acquisition and utilization of digital terrain data.

* Understanding how to quantify the real needs of the user/application for precision rectification and registration of the data and the degree to which analysis results and user end products are affected by errors in registration and rectification.

The areas of advanced technology which were considered seemed to be perceived as somewhat divergent with respect to their prospects for near-term applicability to remote sensing of renewable resources. The emergence of parallel processing systems, capitalizing on the shrinking size and cost of digital computers, was recognized as having great potential for amplifying the rate at which digital imagery can be processed; some systems already exist to do this. General applicability of this form of advanced digital technology may follow from successful research in the direction of

* Memory architecture and management strategies for interfacing parallel processing systems and high-volume, high-dimensional remote sensing data.

* Understanding the theoretical speedup limitations of parallel systems and the concomitant implications for the cost versus benefit tradeoffs involving such systems.

It was further recommended that

* A prototype parallel system using contemporary technology should be assembled to demonstrate the theoretical models and validate performance predictions.

Artificial intelligence, the aim of which is to find ways to make computers perform tasks normally thought of as requiring human intelligence, could eventually lead to automation of the process of obtaining high-level information from pictorial data. Looking at the steps conventionally followed in proceeding from a scene to a description of a scene by way of remote sensing and computer processing, it was observed that artificial intelligence research could contribute to

* Development of scene and sensor models which will allow reduction of raw image data to a form free of incidental variations ("noise" of various forms) without needing local ground truth or ancillary data.

* Development of scene models and analytical mechanisms which will facilitate both the representation and manipulation of information available from a scene (e.g., graph structures and machine-implemented reasoning processes).

But there was some skepticism with respect to the near-term applicability of artificial intelligence research results in remote sensing. Some feel that a more fruitful approach would be to concentrate on facilitating interaction of the human analyst with his data. Still, given that the potential payoff of success in the artificial intelligence domain is very great, near-term progress is hardly a fair criterion for prioritizing fundamental research needs.

In the information extraction sessions, a recurrent notion was "mixtures." More specifically, two significant issues were how to deal with mixtures of dissimilar data in multitype data bases, and how to resolve ambiguities resulting from mixture pixels (often boundaries) in image data. The latter problem stands as a serious barrier to improved spectral classification accuracy and proportion estimation accuracy and is widely recognized as requiring concerted attention. The former represents more an opportunity than a barrier, a source of information about the observed scene which the technology has only begun to exploit. Specific research issues identified include

* Quantifying the effects of mixture pixels on classification and proportion estimation accuracy; finding effective ways to resolve uncertainties arising from the presence of mixture pixels.

* Development of more effective and efficient sampling techniques for classifier training, classifier evaluation, and area/proportion estimation.

* Determining meaningful ways to evaluate and compare alternative methods for proportion estimation.

* Development of effective formalisms for characterizing and differentiating among spatial patterns in complex scenes.

* Development of statistical models and classification methods applicable to data sets with components from greatly different sources.

Appendix 3 contains a more detailed account of the individual presentations and discussions comprising the workshop.

CONCLUSIONS

Overall, the panel of experts did not take issue strongly with any aspect of the Bryant/Guseman report. Quite appropriately, however, there was a strong tendency to focus sharply on basic understanding as opposed to, say, algorithm development. Specifically, the discussions highlighted the need for:

1. Understanding and modeling the physical phenomena which produce deleterious abberations in remote sensing image data.

2. Quantification of user needs for precision in image registration and rectification in order to understand the real value of these operations and impact of residual errors.

3. Understanding the real potential of parallel computing systems for improving the processing efficiency of large remote sensing data sets.

4. Understanding how images capture useful information, how humans extract that information through reasoning processes, and how computers might emulate these processes.

5. Understanding the impact of mixture pixels on scene analysis results and exploration of new approaches for dealing effectively with them.

6. Modeling relationships among diverse data sources and understanding how useful information may be extracted from these relationships.

ACKNOWLEDGEMENTS

This workshop was sponsored in part by the National Aeronautics and Space Administration under NASA Contract NAS9-15466. The assistance of Ms. Shirley M. Davis of LARS and John M. Almon of the Continuing Education Administration, Purdue University, in coordinating physical arrangements for the workshop is gratefully acknowledged.

REFERENCES

1. MacDonald, R.B. and F.G. Hall, "Global Crop Forecasting," Science, Vol. 208, pp. 670-679, May 1980.

2. Barker, G.R., "Resource Information Needs in Industry and the Role of Remote Sensing," Proc. 1981 Machine Processing of Remotely Sensed Data Symposium, Purdue University, June 1981, p. 9 (abstract only).

3. Bryant, J. and L.F. Guseman, Jr., "Basic Research Planning in Mathematical Pattern Recognition and Image Analysis," Department of Mathematics, Texas A&M University, College Station, TX 77843, August 1980.

Appendix 1: WORKSHOP ATTENDEES

Paul E. Anuta
Laboratory for Applications
of Remote Sensing
Purdue University
West Lafayette, IN 47906-1399

Frederick C. Billingsley
Jet Propulsion Laboratory
4800 Oak Grove Dr.
Pasadena, CA 91108

Shirley M. Davis
Laboratory for Applications
of Remote Sensing
Purdue University
West Lafayette, IN 47906-1399

William DiPaolo
Bureau of Land Mgt.
Operations Bldg. 50
Denver, CO 80225

Maoseng Feng
University of Idaho
College of Mines
G-12 Upham Hall
Moscow, ID 83843

Lorrain Giddings
Apartado 63 (INIREB)
Xalapa
Veracruz, Mexico

Lisa Gorgas
Defense Mapping Agency
600 Federal Place
Louisville, KY 40220

Christopher Gunn
University of Kansas
Center for Research
2291 Irving Hill Dr.
Lawrence, KS 66045

Marilyn M. Hixson
Laboratory for Applications
of Remote Sensing
Purdue University
West Lafayette, IN 47906-1399

David A. Landgrebe, Director
Engineering Experiment Station
Purdue University
West Lafayette, IN 47907-0501

Richard S. Latty
Technicolor Graphic Services
Ames Research Center
M/S 242-4
Moffett Field, CA 94035

R. Kent Lennington
1830 NASA Road 1
Mail Code G37
Houston, TX 77059

David Linden
Technicolor Graphic Services
Building 50
Mail Stop D-234
Denver, CO 80225

Fabian Lozano
Apartado 63 (INIREB)
Xalapa
Veracruz, Mexico

Korehiro Maeda
Canada Centre for Remote Sensing
717 Belfast Rd.
Ottawa, Ont. K1A OY7, Canada

Kohtaro Matsumoto
1880 Jindaiji-Suachi
Chofu 182
Tokyo, Japan

Robert B. McEwen
U.S. Geological Survey
519 National Center
Reston, VA 22092

Donald Myhre
2280 NW 21st Place
Gainesville, FL 32605

Terry L. Phillips, Deputy Dir.
Laboratory for Applications
of Remote Sensing
Purdue University
West Lafayette, IN 47906-1399

Ronald Rathnow
Defense Mapping Agency
600 Federal Place
Louisville, KY 40220

Tony Reynolds
Defense Mapping Agency
600 Federal Place
Louisville, KY 40220

Alfredo Rosado
San Antonio Abad 124
1ER Piso
Mexico 8 DF, Mexico

Azriel Rosenfeld
Computer Science Dept.
University of Maryland
College Park, MD 20742

G. Schultink
Michigan State University
Dept. of Resource Dev.
East Lansing, MI 48224

Robert B. Scott
NOAA
3606 T St., NW
Washington, DC 20007

S. Sylvia Shen
Lockheed Corporation
1830 NASA Road One
Houston, TX 77058

Haruhisa Shimoda
Tokai University
1117 Kitakaname
Hiratsuka, Japan

H.J. Siegel
Laboratory for Applications
of Remote Sensing
Purdue University
West Lafayette, IN 47906-1399

David Simonett
Dean, Graduate Division
University of California
Santa Barbara, CA 93106

Mark Sither
Defense Mapping Agency
600 Federal Place
Louisville, KY 40220

Page Spencer
SRA Box 472-F
Anchorage, AK 99507

Philip H. Swain
Laboratory for Applications
of Remote Sensing
Purdue University
West Lafayette, IN 47906-1399

Brian Turner
Pennsylvania State University
228 S. Patterson St.
State College, PA 16801

Lee Werth
2309 Cambridge Ct. N
League City, TX 77573

T.H. Lee Williams
University of Kansas
Center for Research
2291 Irving Hill Dr.
Lawrence, KS 66045

Sheng Xu
3617 Dawson St.
Pittsburgh, PA 15213

Appendix 2: WORKSHOP SCHEDULE

<u>MONDAY, June 22, 1981</u>

7.45 a.m. — Registration

8:00 a.m. - 8:30 a.m. — Opening remarks, charge to the attendees. Philip H. Swain, Purdue University, Workshop Chairman

8:30 a.m. - 11:30 a.m.

<u>Session I: Data Bases and Image Registration</u>

<u>Chairman</u>: David Simonett, University of California, Santa Barbara

<u>Reporter</u>: Paul E. Anuta, Purdue University

DATA BASES FOR REMOTE SENSING. David Simonett

IMAGE PREPROCESSING OPERATIONS. Frederick C. Billingsley

MAP-ORIENTED CONSIDERATIONS. Robert McEwen

12:30 p.m. - 2:30 p.m.

<u>Session II: Advanced Technology</u>

<u>Chairman</u>: Azriel Rosenfeld, University of Maryland

<u>Reporter</u>: Philip H. Swain, Purdue University

ADVANCED DIGITAL SYSTEMS. Howard Jay Siegel

ARTIFICIAL INTELLIGENCE METHODS. Azriel Rosenfeld

3:00 p.m. - 5:30 p.m.

<u>Session III: Information Extraction</u>

<u>Chairman</u>: David A. Landgrebe, Purdue University

<u>Reporter</u>: Richard S. Latty, Technicolor Graphics, Sunnyvale, CA

CLASSIFICATION. Philip H. Swain

CLASSIFIER TRAINING CONSIDERATIONS. R. Kent Lennington

7.30 p.m. - 10.00 p.m. — Group discussions, report formulation.

<u>TUESDAY, June 23, 1981</u>

1:00 p.m. - 4:30 p.m. — Discussion of draft reports.

4:30 p.m. - 5:15 p.m. — Workshop Wrap-Up. Philip H. Swain

Assessment: THE WORKSHOP AND THE FUTURE. Robert B. MacDonald, NASA

Appendix 3: Proceedings of the

WORKSHOP ON KEY ISSUES IN THE ANALYSIS OF REMOTELY SENSED DATA

June 22-23, 1981

SESSION I: DATA BASES AND IMAGE REGISTRATION
Reporter: Mr. Paul E. Anuta

Introduction

The activities of participants in Session I consisted of three overview papers and discussion in the morning session and a discussion the evening of the first day and finally a review discussion the second day. The scope of topics discussed went beyond the title subjects and covered most of the scope of the Bryan/Guseman document (registration, rectification, radiometric correction data structures, and others).

This report is in three parts: (1) an overview of the formal presentations, (2) an account of the discussions that took place, and (3) a statement of the conclusions regarding key issues and changes or additions to the Bryant/Guseman report.

Speaker Presentations

1. DATA BASES FOR REMOTE SENSING -- David Simonett

Dr. Simonett spoke on three questions related to data processing: (1) the general question of rectification, (2) extension to a variety of data sources, and (3) total systems for multiple data sources.

The basic question he posed was: To what extent is high-precision rectification needed by users? He asked: "Is it better to strive for high rectification accuracy or accuracy adequate for purposes at hand," citing the USDA example of tying Landsat results in with their ground-based systems. He pointed out that many microprocessor systems available to users have varying degrees of capability for doing preprocessing. The Bryant/Guseman report stressed <u>absolute</u> accuracy, he said, and he believes this is "largely irrelevant."

The point was also made that whole-frame processing is probably not needed; only small areas are generally processed. "Why do precision processing for all data?" he asked.

Dr. Simonett believes there is a fundamental indeterminacy in the data which limits the ultimate accuracy. He asked what the incremental improvement is which can be obtained by improving registration accuracy. He suggested that this should be a research study.

Examples of registration were discussed: S-192 had interband misregistration. Eastern Maryland area. Needed to know if 1/2 pixel misregistration was a serious problem. In case of radar, could not use topographic highs as controls. Only rivers and long linear features are common to MSS and radar.

With regard to multiple data sources, nominal as well as ordinal data must be considered. The most widely used projection is UTM.

In summary, he listed the following key items:

* The question of how precisely rectification should be done.
* Should we routinely rectify to UTM?
* What are alternative strategies to the whole scene, high-precision registration/rectification approach?

Then slides were shown to illustrate the issues:

1. Regional Analysis for Geology

   Illustrated methods are becoming too automatic and a great deal can be obtained from manual interpretation aids.

2. Land Use Planning and Management

   Percentages of information obtained from remote sensing, given:

   60% certain land use parameters
   30% landscape parameters
   20% socio-economic information

3. Land Use Data Base

   Can improved accuracy be obtained by better sampling rather than by improving registration?

4. Scene and Subscene Statistics (India)

   Are global corrections to whole scene desired?
   Are global scene statistics valid?

5. Area in Australia

6. Uncorrected data

7. Linear stretch

8. Areas predicted as geologically similar

   Fine lines in scene.
   Misregistration would be very serious.

9. Felsic Volcanics

   Serious problem of any misregistration.

10. Question of whether multidate is useful.

11. W. Australia

    Two dates.
    Would precise registration be of benefit?

12. San Francisco Peninsula

    Change in broad areas of interest.
    Would high registration accuracy be needed?

13. W. Australia

    Enhancements vs. multidate.
    Ratios, principal components, and Band 4/PC1.
    Abundant geological information.

"We need to define the degree to which problems presented in the Bryant/Guseman report are crucial to success of remote sensing goals," Dr. Simonett said.

2. IMAGE PROCESSING OPERATIONS -- Frederick C. Billingsley

Dr. Billingsley discussed preprocessing problems generally, following the Bryant/Guseman report. He presented a set of slides and overheads containing key items for consideration and also some interesting research results. The overheads are reproduced on the following pages. Three slides were presented first, containing results of studies on effects of noise and misregistration on classification accuracy. It was pointed out that misregistration causes significant loss of accuracy for crop fields of typical midwest size.

FUNDAMENTAL RESEARCH WORKSHOP

PREPROCESSING TOPICS

- GEOMETRIC
- RADIOMETRIC
- SPATIAL FREQUENCY
- GEOREFERENCING AND CATALOGUING

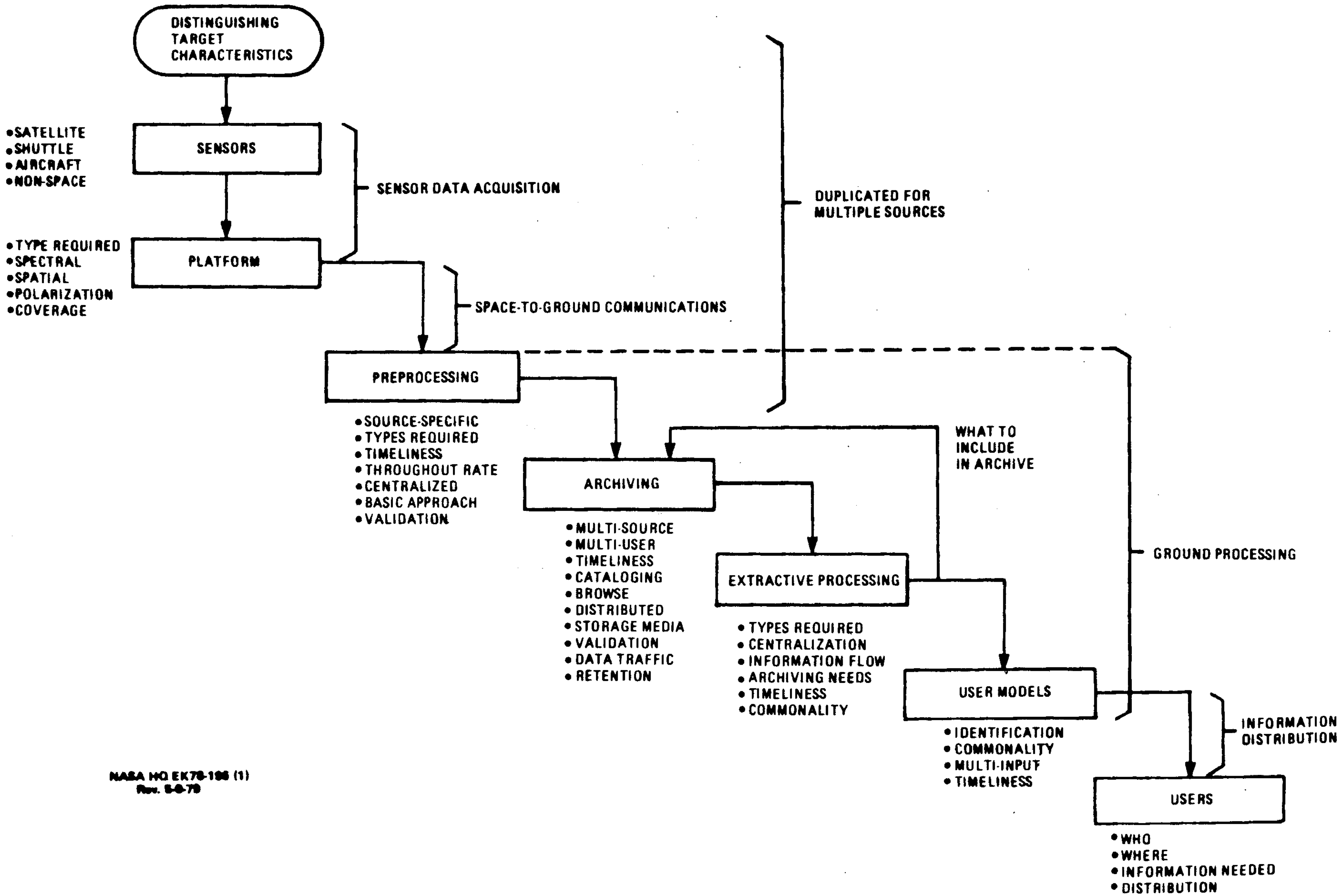
PREPROCESSING FUNCTIONS

TO MAKE THE DATA

- CORRECT
- COMPATIBLE
- AVAILABLE

FCB-6/19/81

TECHNOLOGICAL COMPONENTS AND ISSUES OF TOTAL INFORMATION SYSTEM STRUCTURE



NASA HQ EK78-196 (1)
Rev. 5-8-78

FUNDAMENTAL RESEARCH WORKSHOP

ASSUMPTIONS

- DATA ACQUIRED, STORED, USED DIGITALLY
- IN SURVEY (LANDSAT) MODE, ONLY A SMALL FRACTION OF THE DATA IS USED DIGITALLY
- SOME PROCESSING (E.G., MAP PROJECTION) DONE ON RETRIEVAL

FCB-6/19/81

FUNDAMENTAL RESEARCH WORKSHOP

CAUSES OF DISTORTION

	INTERNAL	EXTERNAL
RADIOMETRIC	SENSOR LINEARITY DIFFERENCES BETWEEN SPECTRAL RESPONSES FOR MULTIPLE DETECTORS	ATMOSPHERE HAZE BIDIRECTIONAL REFL ILLUMINATION ATMOSPHERE SPECTRAL
GEOMETRIC	SCAN NONLINEARITY JITTER PLATFORM MOTION OPTICS AND FILM	PROJECTION RELIEF DISPLACEMENT
SPATIAL FREQUENCY	APERTURE ELECTRONICS OPTICS INTERPOLATIONS	ATMOSPHERE PSF

FCB-6/19/81

FUNDAMENTAL RESEARCH WORKSHOP

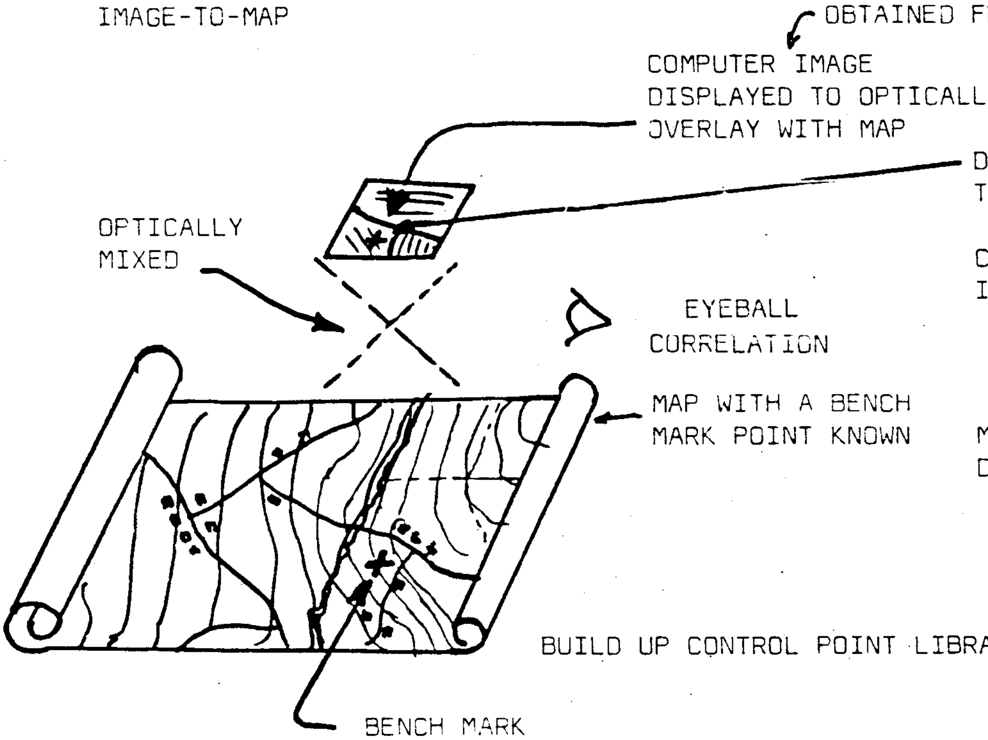
CONTROL AND CORRESPONDENCE

• CONTROL POINT TYPES AND WHEN TO USE	I
• CORRELATION TECHNIQUES FOR CONTROL POINTS	I
• CONTROL POINT LOCATION WITHOUT CORRELATION	I
• HOW CAN DIGITAL ELEVATION MODEL OR DIGITAL LINE GRAPHS BE USED FOR CONTROL?	II
• HOW TO USE CONTROL POINTS WHICH VARY SEASONALLY	II
• HOW MUCH CONTROL IS NECESSARY	II
• HOW TO HANDLE CONTROL FOR DISPARATE DATA	III
• IS CONTROL METHOD THE SAME FOR RELATIVE AND ABSOLUTE?	III

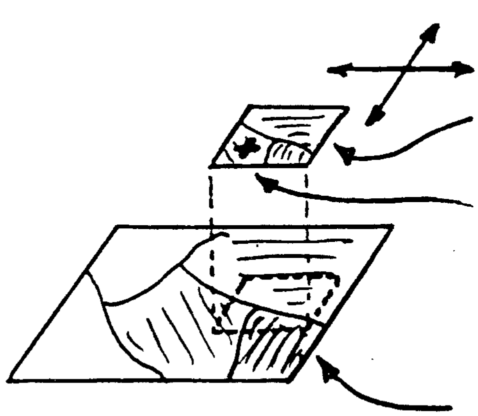
FCB-6/19/81

GROUND CONTROL POINT OPERATIONS

IMAGE-TO-MAP



CORRELATION FOR CONTROL POINT LOCATION BY CONVOLUTION



FUNDAMENTAL RESEARCH WORKSHOP

RESAMPLING

- (HOW) DOES INTERPOLATION AFFECT INTERPRETATION? I
- (HOW) TO WARP INFERENCES AFTER INTERPRETATION? I
- IN WHAT ORDER TO USE N.N. WARPING/ RESAMPLING? I
- CAN SYSTEMS BE DESIGNED TO AVOID WARPING? II
- HOW TO INTERPOLATE WARPING MODEL BETWEEN CONTROL POINTS?

FCB-6/19/81

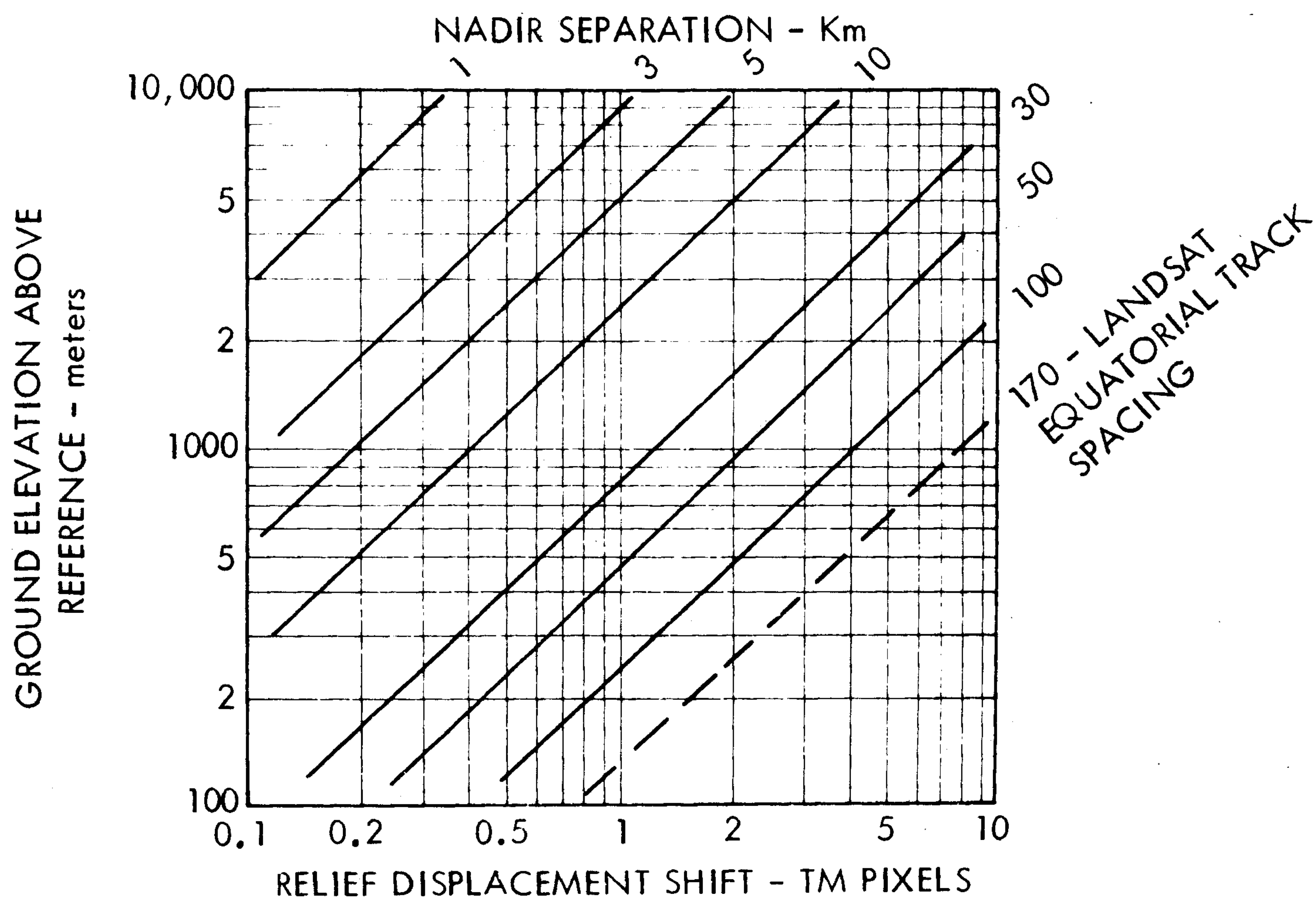
FUNDAMENTAL RESEARCH WORKSHOP

PLATFORM MODELING

- PARAMETRIC SENSOR GEOMETRIC MODEL I
  - EFFECTS OF RESOLUTION
  - EXTERNAL MODEL TO INCLUDE RELIEF DISPLACEMENT
- (HOW WELL) CAN KALMAN FILTERING IMPROVE ATTITUDE MODEL? I
- ON-BOARD DATA BASE FOR REAL-TIME LOCATION II
- REAL-TIME RECTIFICATION ON BOARD? III

FCB-6/19/81

RELIEF DISPLACEMENT EFFECTS



FUNDAMENTAL RESEARCH WORKSHOP

RADIOMETRIC ISSUES

- WHAT IS UTILITY OF ABSOLUTE CALIBRATION?
- ATMOSPHERE CONTRIBUTES HAZE, SPECTRAL SIGNATURE, PSF
- NON-MEASURED FACTORS (E.G., SOIL MOISTURE, RELIEF) CAUSE SPECTRAL DISTORTIONS
- CHANGES IN VIEWING AND ILLUMINATION ANGLES CAUSE SPECTRAL DISTORTIONS
- SENSOR/SCENE POLARIZATIONS WILL INTERACT

FCB-6/19/81

FUNDAMENTAL RESEARCH WORKSHOP

SPATIAL FREQUENCY ISSUES

- INSTRUMENT PSF GENERALLY UNDERSTOOD, BUT CORRECTIONS NEED DISSEMINATION

- ATMOSPHERE PSF UNKNOWN

- INTERPOLATION FREQUENCY RESPONSES UNDERSTOOD BUT NEED REITERATION

FCB-6/19/81

FUNDAMENTAL RESEARCH WORKSHOP

ERROR ESTIMATIONS

• DETERMINE/DEFINE ERROR MEASURES	I
• GROUND LOCATION ACCURACY MAPS	I
• TEMPORAL OVERLAY ACCURACY	I
• RADIOMETRIC ERROR BUDGET	I
• RADIOMETRIC MOSAIC SEAM PROBLEM	II
• RELATIVE VALUES/DIFFICULTIES OF ABSOLUTE AND RELATIVE ACCURACIES	II

FCB-6/19/81

FUNDAMENTAL RESEARCH WORKSHOP

ANCILLARY SENSING

- ATMOSPHERE FOR DATA CORRECTION
- PRECISION ATTITUDE SENSING
  - GROUND LOCATION
  - RECONSTRUCT LINE ARRAY IMAGES
- IN SITU STEREO FOR RELIEF DISPLACEMENT CORRECTION
- FOR "SMART" SENSING - E.G., LOOKAHEAD
- EVANS' MIRRORS FOR BENCHMARK LOCATION

FCB-6/19/81

FUNDAMENTAL RESEARCH WORKSHOP

IMPLICATIONS TO SENSOR SYSTEMS

- FUTURE STRESS ON EASE OF DATA PROCESSING, GREATER INSTRUMENT ACCURACY
- SENSOR GROUPING WILL LEAD TO PLATFORM APPROACH - GROUPING ON ORBIT, TIME OF DAY
- GEOGRAPHIC REGISTRATION REQUIREMENTS LEAD TO NEED FOR BETTER POSITIONING AND POINTING ACCURACY
- TIMELINESS AND DIRECT RECEPTION LEAD TO DESIRABILITY OF SOME ON-BOARD PROCESSING
- INCREASED RESOLUTION, PARAMETERS MEASURED WILL LEAD TO NEED FOR DATA COMPRESSION/AVOIDANCE - POINTABLE SENSORS WITH SELECTABLE RESOLUTION
- GROUND PROCESSING CAPABILITIES MUST INCREASE - USE OF SPECIAL PURPOSE AND PARALLEL PROCESSING INDICATED

FCB-6/19/81

EASCON, 1978

FUNDAMENTAL RESEARCH WORKSHOP

GEOMETRIC ERROR TABLE

CAUSES	MODEL	EXPECTED ERRORS (TM)
	SENSOR/SPACECRAFT DIST'S	NYQUIST IS SMALL
ANTENNA, SOLAR PANEL MOTION	JITTER	SINUSOIDAL - $\approx$ 10 PIXELS/CYCLE
ATTITUDE CONTROL SYSTEM	ATTITUDE DRIFT	SLOW, SYSTEMATIC - $\approx$ SEVERAL CYCLE/ FRAME POTENTIAL JITTER DURING CONTROL - SEVERAL PIXELS
SCAN LINE CORRECTOR	SCAN CORR (TM)	SEVERAL CYCLES/LINE
SPRING MOUNT	MIRROR SCAN	$\approx$ 1 CYCLE/LINE
	EXTERNAL DIST'S	NYQUIST IS LARGE
ATTITUDE CONTROL ACCURACY	ATTITUDE PROJECTION	MANY FRAMES/CYCLE, UNLESS ATTITUDE JITTER
Δ ALTITUDE, GEOID	ROUND EARTH	MANY FRAMES/CYCLE
EARTH ROTATION	SKEW	CALCULATABLE TO FRACTIONAL PIXEL
EPHEM CONTROL, TIMING	S/C POSITION (EPHEM)	SEVERAL KM
EPHEM CONTROL	ALTITUDE	GLOBAL
	SCENE DIST'S	
TOPOGRAPHY, LACK OF S/C COINCIDENCE	RELIEF DISPLACEMENT	COULD BE PER PIXEL
BASIC SURVEY, GEOID	GCP INACCURACY MAP LOCATION	FEW TO MANY PIXELS TO NO MAP
SEASONAL, LACK OF GOOD POINTS IN REQ'D LOCATION, CORRELATION TECHNIQUE	ABILITY TO CORRELATE	PARTIAL TO MANY PIXELS
	RCP INACCURACY	
SEASONAL, LACK OF REQ'D POINTS, CORRELATION TECHNIQUE	ABILITY TO CORRELATE	PARTIAL TO MANY PIXELS
DIFFERENCE BETWEEN WARPING MODEL AND REALITY	SCENE WARPING MODEL	PARTIAL TO MANY PIXELS
ALGORITHM SELECTION	NEAREST-NEIGHBOR WARPING	PARTIAL PIXEL

FCB-6/19/81

3. MAP-ORIENTED CONSIDERATIONS -- Robert McEwen

Dr. McEwen opened with a slide of a French news article of 1874 stating that photographic surveys from balloons were impractical due to high cost. He then turned to the Cartography vs. Remote Sensing dichotomy:

There is a feeling towards not using cartographic terms. USGS could not call the Florida mosaic a map. What do cartographers do? They endeavor to do categorization and spend a great deal of time deciding what should be mapped. Categories must have meaning to someone.

Registration -- in the domain of one sensor. If we move outside to other sensors, then we need to do rectification.

Digital Cartography -- (This is not automatic cartography.) The ability to convert map data into a computer-readable environment is a key requirement. How do you tell a computer what is next to what? Topologic relationships are important. It will be possible to combine remote sensing and map information in the computer to form a powerful data base. Example: Digital Terrain: USGS digitizing 7-1/2 minute quadrangles. Maps have 7m vertical accuracy, 30m pixel spacing. These data can help in Lambertian considerations. Other planimetric categories are being digitized, e.g., public land boundaries, stream courses. Logic is needed for closing polygons. "The Digital Cartographic Revolution will have a profound impact on remote sensing."

Models -- Geometric models lousy. Always operationally changing our modus operandi. Many models studied.

Issues -- Settle down to doing business certain ways. Have to deal with a lot of ground control points. How many are needed? 12 - 180 to check accuracy. NASA uses 40 for current correction in the master data processor.

Photogrammetry -- Put photos together in blocks rather than try to get control for each photo. Blocks joined mathematically. Boundary and cantilever problems. Better to determine and control the attitude of platform. Inertial systems are getting good enough to eliminate need for control. Photogrammetric surveys without control are possible.

Rule of thumb: 10-to-1 ratio of total error to individual error.

Issue of Design -- What is user perception of output of system? Need to do more in design of the output.

Publication -- We are moving into electronic media. Instant throwaway maps will become more common.

Overview Discussion

* It was stated that renewable resources applications do need cartographic accuracy, so this was a relevant subject for research.

* On the Registration/Rectification Dichotomy: The difficulty in achieving given accuracies needs to be established. Provide new materials to the user. Let the user determine the need and cost to achieve a given level of registration/rectification accuracy.

* Dr. Haralick suggested an intermediate product with control. It was pointed out to him that this is the format of the EDC "A" tape. Dr. Landgrebe pointed out that we need to worry more about the basic research questions and not about procedural questions.

* Dr. Haralick also questioned the need of subpixel registration, asking: "Why not find classifiers which can perform well under misregistration?"

* We need to determine the cause of bias in extending class decision to field boundaries.

* We need to look at historical boundary location to predict current locations.

Key Issues From Data Bases and Image Registration

* Evaluation of the benefit of improved registration/rectification must be done.

* Investigate other strategies for registration/rectification rather than total processing.

* Must include spatial frequency effects analysis in with previously cited radiometric and geometric considerations, and also include atmospheric point-spread function.

* Must continue to pursue better platform control and modeling to eliminate need for registration/rectification.

* Must address the issue of what are the desirable cartographic products from remote sensors.

SESSION II: ADVANCED TECHNOLOGY
Reporter: Dr. Philip H. Swain

Introduction

The purpose of the Advanced Technology session was to assess the potential roles which selected subfields of computer science might play if a focused effort were made to apply them to remote sensing of renewable resources. The formal speakers, both of whom are very well known in their own areas of expertise and well acquainted with the general remote sensing problem, discussed the application of advanced digital systems and methods of artificial intelligence.

Speaker Presentations

1. ADVANCED DIGITAL SYSTEMS -- Howard Jay Siegel

Dr. Siegel introduced the basic concepts and terminology necessary to discuss parallel computing methods and systems and gave several examples of parallel algorithms applicable to remote sensing image processing. The examples included maximum likelihood classification, image smoothing, histogramming, and two-dimensional Fast Fourier Transforms. He pointed out that some parallel machines have already appeared on which these algorithms can be (in some cases, have been) implemented. Familiar examples include STARAN and Illiac IV. These computing systems were developed to do general parallel processing tasks. Relatively little research has been done to exploit the potential of parallel processing specifically for multivariate image processing.

Dr. Siegel described some testbed systems being developed which could provide opportunities to assess this potential. Compared to existing systems, the embryonic systems will have considerable architectural flexibility and have far better facilities for program development and testing.

From Dr. Siegel's presentation and the associated discussion, the following research recommendations emerged:

1. Investigate and evaluate in quantitative terms the potential benefits which could be derived by applying parallel processing to remote sensing image processing tasks.

2. Develop memory management strategies for interfacing parallel processing systems and high-volume, high-dimensional remote sensing data. Memory management and associated input/output represent the most serious bottleneck in parallel preocessing of large-scale remote sensing imagery.

3. Carry out a realistic cost/benefit study on the use of parallel systems in remote sensing applications using (a) off-the-shelf component characteristics, and (b) foreseeable digital computer technology.

4. Develop methods for predicting theoretical speedup limitations for parallel implementations of remote sensing tasks.

5. Develop implementation procedures which facilitate optimal implementations of remote sensing processing algorithms.

6. Develop a prototype single-instruction-stream multiple-data-stream (SIMD) system (e.g., an array of microprocessors) to demonstrate the implementation of theoretical models and validate performance predictions.

These recommendations are in concert with, but somewhat more specific than, the recommendations which appoeared in the Bryant/Guseman reference report.

2. METHODS OF ARTIFICIAL INTELLIGENCE -- Azriel Rosenfeld

Confessing to not being in the business specifically of making things easier for remote sensing, Dr. Rosenfeld described his research interest as making computers do intelligent things (not necessarily by practical means). Given the volume of data to be analyzed and the relatively slow, inconsistent and labor-intensive methods now widely used for renewable resources applications of remote sensing, Rosenfeld's interest is very relevant to the topic at hand. Fundamental research along these lines may lead to important breakthroughs in both the effectiveness and efficiency of remote sensing data analysis.

Rosenfeld's presentation and the ensuing discussions raised the following research issues which are or could be addressed by the methods/concepts/philosophies of artificial intelligence (refer to the figure on the next page):

1. Develop scene-dependent models for image correction, to determine and remove systematic distortion, which permit automatic determination of corrections from the image itself.

2. Develop methods for deriving intrinsic images (images devoid of incidental, noninformation-bearing variations) independent of ancillary data (e.g., remove effects of terrain relief from spectral response data).

3. Develop region-level models for scene segmentation and abstraction to derive symbolic images for manipulation and analysis.

4. Investigate the use of hierarchical graph structures (pyramids, quadtrees, etc.) for representing scene information content.

5. Develop analytical methods for producing informative descriptions from symbolic representations, e.g., reasoning processes such as are embodied in "expert systems."

IMAGE ANALYSIS PARADIGM (Due to A. Rosenfeld)

Data	Process	Knowledge	Issues
Scene			
↓	Sensing		
Image			
↓	Correction	Models of the sensing process	Models for the scene? Most processes are to some extent scene dependent.
Corrected Image			
↓	Disambiguation	Elevation or slope maps, etc.	Need "bootstrap" process not requiring ancillary data
"Intrinsic Images"			
↓	Segmentation	Scene models: pixel level	Region-level models? Do useful models exist?
Symbolic Images, Regions			
↓	Representation		Hierarchical array structures (pyramids, quadtrees as intermediates?)
Graphs (Hierarchical)			
↓	Identification	Maps	Constraints? Bases, mechanisms for reasoning?
Description			

In the discussion, Rosenfeld was asked whether, in the perspective of the next 5 - 10 years, it really makes sense to put a lot of emphasis on development of artificial intelligence methods as a substitute for user/analyst interaction with the data and processing results. The implication was that the artificial intelligence approach may be a very long way from being developed to the level of practical application. Rosenfeld pointed out, however, that some rather promising results are already being achieved in applying these techniques to mineral prospecting via remote sensing imagery and other uses of aerial interpretation both here and abroad.

SESSION III: INFORMATION EXTRACTION (Object Scene Inference)
Reporter: Dr. Richard S. Latty

Introduction

The purpose of this documentation is to augment and complement the content of the Bryant/Guseman report directed at identifying, formalizing, and prioritizing the current issues in digital image analysis. The manner in which this is conducted is through the compilation of material presented and the ensuing discussions in the area of object scene inference. This material is organized by speakers since each represented a sufficiently distinct area and frame of reference pertaining to object scene inference. The background material presented is summarized to provide orientation.

Speaker Presentations

1. SYSTEMS CONCEPT OF INFORMATION EXTRACTION -- Philip H. Swain

Information extraction in remote sensing, from a total systems standpoint, is the process of transforming the actual physical scene into a body of information. This systems level process involves many related but distinct processes. These usually involve: Distributions of electromagnetic energy incident on the scene; object composition and orientation relative to the source of illumination; the consequent reflectance, or emittance; atmospheric modulation and attenuation; photon reception; amplification (which may involve a transformation, e.g., log transform); quantization; recording and/or telemetry; satellite relay and/or ground reception; conversion; calibration (band-to-band, within band); geometric and response level rectification or adjustment; generation of a discriminant function or decision rule (in a concrete form), or some other activity intended to convert some representation of measured irradiance values into some "meaningful" information; and higher order information generation through modeling or data base integration and manipulation.

While the concern throughout this presentation is with the extraction subprocess, the success attained in meeting the objectives of the analysis is dependent on the degree to which properties of interest are contained in the scene, and to what degree these properties are preserved in the data. What is not intrinsically contained in the scene cannot be contained in the data obtained from the scene. Secondly, the desired information must be retained in the data from the scene. The measurable properties of the data must be consistently dependent on the properties of the scene which are of interest with respect to the objectives of the processing task. These properties may be measurable as patterns in the spectral domain (multispectral data), the temporal domain (multidate data), the spatial domain (as in textural computations or spatial associations), or some combination of these and other domains.

The activity of information extraction must presume that the desired information is contained in the properties of the scene, is retained in the data, and is of sufficient impact on the outcome of the decision-making process to warrant the information extraction activity. These can be considered as real-variable, continuous constraints on the warranted level of expenditure for the information extraction activity. The remaining requisite for information extraction is to identify how the information is contained in the data; that is, what are the information-bearing characteristics (the "features") of the imaged scene? Is the information represented in patterns in the spectral, temporal, spatial, or some other domain? Once the information-bearing features are identified, the problem becomes one of selecting a method of information extraction. Candidate procedures can be evaluated relative to an increasingly critical sequence of selection criteria. First, the candidates must be admissible; they must provide, to some degree, the desired information. Secondly, they must be feasible procedures; they must not require unavailable resources nor place unmanageable demands on the available resources. Lastly, the selected procedure would ideally be "optimal"; that is, the ratio of the value of the information extracted to the cost of the information obtained would be greatest of all candidate techniques.

Candidate techniques for extracting the information from the data include:

1. Various data transformation activities conducted to render the data more compatible with the particular extraction technique to be used.

   For change detection, this might involve band ratios.

   For multi-univariate classifiers, this might involve a principal components analysis, canonical analysis, or some other dimension reduction transformation.

2. "Per-point" classification algorithms.

   Parallelpiped classifier.

   Minimum Euclidean distance classifier.

   Gaussian maximum likelihood classifer (assumes a covariance which varies with respect to cover class).

3. "Sample" classifiers.

   Classification decision rules based on statistical distance measures between multivariate Gaussian distributions.

   Employ either fixed or variable pixel grouping size; variable pixel grouping size requires image partitioning to precede the classification.

4. Textural measures computed over a moving window.

   Vary widely in specific form and computation, but can be categorized as being either a specified computational transform (e.g., entropy, second angular moment,...), or least squares estimation of descriptive coefficients (e.g., facet model, Fourier series,...).

5. Algorithms which examine the neighborhood of pixels and employ measures of class frequency, or examine the relative evaluated probability density function associated with the neighborhood, fall in the category of contextual classifiers.

6. Algorithms which arrange the decision logic in a sequential or decision tree approach provide a means of employing widely different forms of data in the information extraction process.

Information extraction employing remotely sensed data does not cease after the process of classification but continues through the integration with other data types. Through modeling and other inference-making processes, the information extraction process encompasses a much broader range of operations.

Research Issues

1. Accurate statistical models of the data are needed for data of widely different data sources (from different scanners, linear arrays, radar, passive microwave, geophysical data, scatterometers, map bases, terrain data, soils data, meterological data, geodetic data,...).

2. More flexible multistage analysis techniques are needed. This is dependent on a more thorough understanding of the relationships between data type and information type in a multi-node or multi-level decision tree process approach to information extraction.

3. The need exists for a more thorough understanding of how to formalize spatial relations in digital data and how these relations correspond to the desired information.

4. How is mathematical formalization of spatial relations to be assessed relative to other formalizations or computations in a manner which is meaningful in terms of accuracy and precision in the information extraction process?

5. A thorough understanding is needed of what determines the extent of the region for which spatial relations consistently represent information about the scene components. A means of determining the region size for different scenes also is needed.

2. PROPORTION ESTIMATION -- R. Kent Lennington

Proportion estimation is concerned with obtaining an estimate of the relative area occupied by each constituent ground cover in a given scene. Proportion estimations can be obtained by any of several analysis procedures embodied in the techniques employed by remote sensing. The task of training the classifier provides an estimate of the areas occupied by each cover class if the samples are selected at random, are sufficiently numerous to provide stable estimates (i.e., small variances associated with each estimated proportion), and each sample unit is accurately identified. Problems arise in the implementation of the training techniques due to: inabilities to correctly identify the sample unit in terms of the classes contained in the scene, inabilities to locate the sample unit in terms of the classes contained in the scene, inabilities to locate the sample unit accurately in the reference data (registration error); also the spatial extent of the sample unit may not be compatible with the spatial extent of some of the classes in the scene (error due to spatial resolution). Other problems arise due to the cost of acquiring a large sample size.

While classification and proportion estimation place similar constraints on the training procedure, the relationship between classification and proportion estimation is less clear. Classification provides a statistically biased proportion estimate due to the inequality of the errors of omission and commission, which arise from inequality of the spectral covariances between the classes in the scene. Furthermore, if classification is to be warranted in addition to the estimate obtained from the training procedure, the number of samples used in training the classifier should be relatively small. Hence the classification is either likely to be inaccurate or constitute an unwarranted additional expense.

Attempts to reduce the bias introduced by classification have been made by using the classification to stratify the scene for the purpose of sampling effort allocation. The process is generally regarded as wasteful and equally good estimates are provided by increasing the initial sampling effort.

Other approaches employed in obtaining the proportion estimate include employing a small and simple training sample. Estimating the classifier and the omission/commission matrix is conducted through a jackknife approach. The proportion estimates are then adjusted through the estimated omission and commission frequencies. Clustering the entire scene, to stratify the scene, also has been employed as a means of sample effort allocation. These techniques have met with similar problems encountered in the aforementioned approaches.

The need to accurately label the sample units is in direct conflict with the need for the sample units to be representative of the population of which the sample units are members. In order to be representative, the sample must include atypical (in a spectral sense) as well as typical sample units. In order to achieve this representation, the sampling should be objective (e.g., systematic random sample). Typical samples are generally labeled with a high degree of accuracy. Atypical samples are labeled with lower accuracy. Methods of resolving this problem have employed attempts to estimate the error frequency associated with labeling and adjust for this frequency in the classifier design or proportion estimate.

Another, more promising, approach is to adopt a method wherein the sample includes typical and atypical sample units but only the typical sample units are employed in the labeling process. A mixture model is employed to formalize the problem and assist in designing the approach employed to resolve the problem.

The component densities of the mixture are assumed to be multivariate-normal and hence each component density is distinguished from other components in the mixture through a Gaussian, iterative estimation technique employing a sequence of split/combine processes. This provides an objective means of associating the atypical and typical samples (provided the multivariate-normal assumption is satisfied and that for every mode there occur typical samples). The proportion estimate can then be obtained through a summation of the components associated with each class in the scene.

Research Issues

Implicit to the area of proportion estimation based on spectral data are numerous problems associated with the relationship between patterns in the spectral domain and each class and mixtures of classes contained in the scene. Problems which arise pertaining to this relationship are:

1. What is the influence of boundary pixels on the accuracy of proportion estimation?

2. How can boundary pixels best be accommodated in the context of proportion estimation?

3. If area estimation is being conducted for crop production estimation, is there information germane to this objective in the evaluated discriminant functions or the relative location of "field-center" pixels and boundary pixels in the spectral space?

4. What types of information that may be available should be employed in systematizing the sampling allocation in order to obtain a more representative and effective sample?

5. How are various proportion estimation techniques to be compared in a meaningful and consistent manner?

3. SPATIAL PATTERNS -- Robert M. Haralick
(Based on material submitted after the Workshop)

The limitations encountered in information extraction from digital images through employment of per-point classifiers, or analysis approaches which examine the gray level value variation over small local regions, become more pronounced for more complex scenes and imagery obtained at higher spatial resolutions. Overcoming these limitations will be dependent on advances in use of spatial patterns in the image. To facilitate these efforts, an exact and comprehensive language of spatial patterns is needed. Only with such a language can the approach be formalized and structured into algorithms compatible with machine processing.

Spatial patterns are a function of the organization of physical objects in the three-dimensional spatial world and are rendered apparent through the presence of a spectral difference among the objects arranged in the 3-D space. These are the "ground spatial patterns." These spatial patterns are modified and transferred to the image, depending on the geometric relations between the source of illumination, the reflecting surfaces of the objects in the scene, the point and orientation at which each fraction of the scene is imaged, and the spatial extent of each fraction. The resulting spatial pattern is the "image spatial pattern." Therefore, an association exists between image spatial pattern and ground spatial pattern. Such an association could be formalized through the conditional probability:

$$P(I|G) = P(\text{image spatial pattern} \mid \text{ground spatial pattern})$$

To actually implement this formalization in a decision rule, we need a means of parametrically defining each component of the conditional probability. The parametric representation is dependent on the language of spatial patterns. This language is the means of describing the data structure which constitutes the image spatial pattern. To be compatible with the idea of parametric representation, there must be one data structure which serves as a good, or central, representation of each generic kind of spatial pattern. Generic kind connotes some level of dissimilarity from all other existing data structures. The level of dissimilarity needs to be measurable through some real number representation which can be determined through a comparison of the representative data structures.

Research Issues

1. There is a need then to evolve a concise language of spatial patterns.

2. The appropriate data structures with which to represent each generic kind of spatial pattern need to be determined.

3. A dissimilarity function is needed to represent the separability of these data structures.