DEVELOPMENT OF ALTERNATIVE DATA ANALYSIS TECHNIQUES FOR IMPROVING THE ACCURACY AND SPECIFICITY OF NATURAL RESOURCE INVENTORIES MADE WITH DIGITAL REMOTE SENSING DATA

Investigators: Dr. Thomas M. Lillesand
Douglas E. Meisner
Remote Sensing Laboratory
Institute of Agriculture
University of Minnesota
St. Paul, Minnesota

INTRODUCTION

General application of digital remote sensing technology has gone through considerable change in the recent past. In spite of growing user interest in applications of digital classification for thematic mapping, the private sector has shifted its focus somewhat away from this activity. Attention is being concentrated more on hardware development and provision of digital image enhancement services, rather than on classification per se. This shift is partly due to the increasing sophistication of end users as the "gee whiz" allure of the technology wears off, a smaller core of more serious users is emerging. This is a healthy development. Another reason for the shift, however, is the difficulty and cost inherent in performing digital classification services in a non-research environment. The need for active involvement of field personnel, coupled with the problems of training those personnel in the generally complex procedures, has frustrated many analysis efforts. At the same time, as accessible as remote job entry terminals have become, line printer output simply does not afford the graphic interaction required for most sophisticated classification procedures. In this context, one can understand the shift toward image enhancement procedures; these activities are greatly simplified by the clear and convenient separation of digital processing efforts from subsequent field analysis. Yet, as training courses (e.g. at the three NASA regional centers and at the EROS Data Center) continue to nurture interest in the use of digital classification techniques, the need for serving those users will increase. This is particularly critical as applications spread to the state and regional level, since these users are less able to obtain specialized equipment in-house.

The work begun under this grant is an investigation into the ways to improve the involvement of state and local user personnel in the digital image analysis process. The intent is to isolate those elements of the analysis process which require extensive involvement by field personnel and to provide means for performing those activities apart from a computer facility. In this way, the analysis procedure can be converted from a centralized activity focused on a computer facility to a distributed activity in which users can interact with the data at the field office level (or indeed in the field itself). This concept is illustrated in Figure 1. If successfully implemented, the distributed approach would offer these advantages:

1. Provide more efficient use of computer resources, by reducing the digital image processing effort to a highly standardized procedure which can be run in a batch mode.
2. Provide more economical use of field personnel, by eliminating the expense of transporting them to the analysis facility and lodging them while there. "Day to day" work could then be handled while

A second preliminary analysis involved forest typing with LANDSAT data on a College of Forestry test site for which detailed ground data and high altitude color infrared photography were available. In this case, training sets were delineated on high altitude photography and digitally transformed into LANDSAT scene coordinates. Unfortunately, the second order least squares polynominal fit used in the transformation was inadequate for relating the photographic base to the LANDSAT data. Due to the spatial complexity of the study area, even the slight displacements resulting from the transformation process affected the training statistics excessively. An interesting sidelight was the fact that the RMS errors computed in the polynomial fitting process were very low, suggesting a higher accuracy than was actually realized on a training area by area basis.

Subsequently, the training process was performed by visually relating the high altitude images and the hard copy LANDSAT enlargement in a Zoom Transfer Scope. This approach enabled features in the LANDSAT enlargement to be identified with much more certainty and led to successful training. This process was found to be impossible using line printer output.

3. Procurement of digital display equipment. In-house funds have been used to acquire a stand-alone microcomputer-based image display system which will be used on this project. Installation of this equipment was completed in November. The system consists of a Spatial Data Eye-Com video digitizer and display, a DEC LSI-11 microcomputer with Fortran compiler, a DeAnza Visacom color display system, color

vs. Dicomed output, log vs. linear recording, various contrast stretch techniques, and different film and print paper types.

The use of film recorder images as a base for supervised training was tested in several pilot studies. This began with several student projects in which the color data were used for visual analysis but actual training set delineation was performed on line printer image output. The availability of the color data for interpretation of spectral characteristics enhanced the analysis process greatly, but the training set delineation was still a cumbersome task.

Two other preliminary analyses have been performed. The first involved classification of wetland types on a photograph which had been digitized using a scanning densitometer.[1] Training sites were located on the color image base and digitally transformed (using a polynomial transformation model) into scene coordinates. Because the encoding was done by hand, the process was still cumbersome. However, this was seen as a simulation of a coordinate digitizer-based procedure to be developed in the near future.

The polygon processing software was also used in this study to mask the image data set, in order to restrict the classification to selected sub-scenes (in this case, wetland areas). This permitted successful classification of within-wetland species, a task which had been previously impossible because of spectral confusion between some wetland and non-wetland classes.

See Section A of this report.

PROGRESS TO DATE

In the past six months of work on this project, the following four general areas of activity have been pursued:

1. Development of general image processing software on the University of Minnesota computer system (Control Data Cyber models 172 and 74). This software enables us to read CCT's in several formats, including the old and new LANDSAT tape formats, scanning microdensitometer tapes obtained from the University of Wisconsin, and tapes generated by the Data Analysis Laboratory of the EROS Data Center. Programs to support supervised training and optimized maximum likelihood classification have been written and used in this and other research projects. Interfaces to the U of M Dicomed color image recorder have been developed, both to display contrast enhanced original image data and to record classification output results.

2. Initial investigation into the use of color hardcopy image data as a primary medium in supervised training procedures. The procedure we employ is to generate digitally enlarged and contrast enhanced hard copy color prints of image data within study areas. These prints are then used to analyze visually (rather than statistically) the spectral properties of the data before supervised training sites are selected on the hard copy print. This technique of supervised training on hard copy images is perceived as a "first look" at a distributed analysis approach.

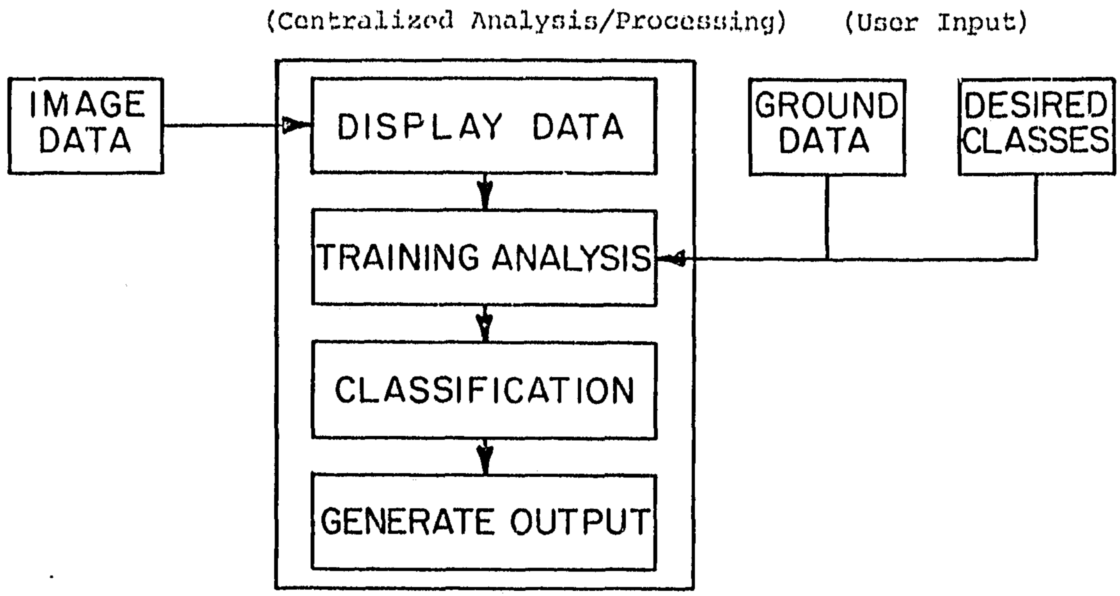
   The investigation began with a basic examination of image recording techniques, including comparisons of video displays working on the classification project. State employees in Minnesota have stressed the importance of this point. By making agency involvement more economical in various ways, a larger number of personnel may participate for longer periods of time.

3. Reduce the difficulties of working with untrained personnel by eliminating their direct involvement with a computer system and somewhat standardizing the analysis procedure they employ. This would remove the user's exposure to the computer operating system and the inevitable system downtime, which are generally frustrating to the new user.

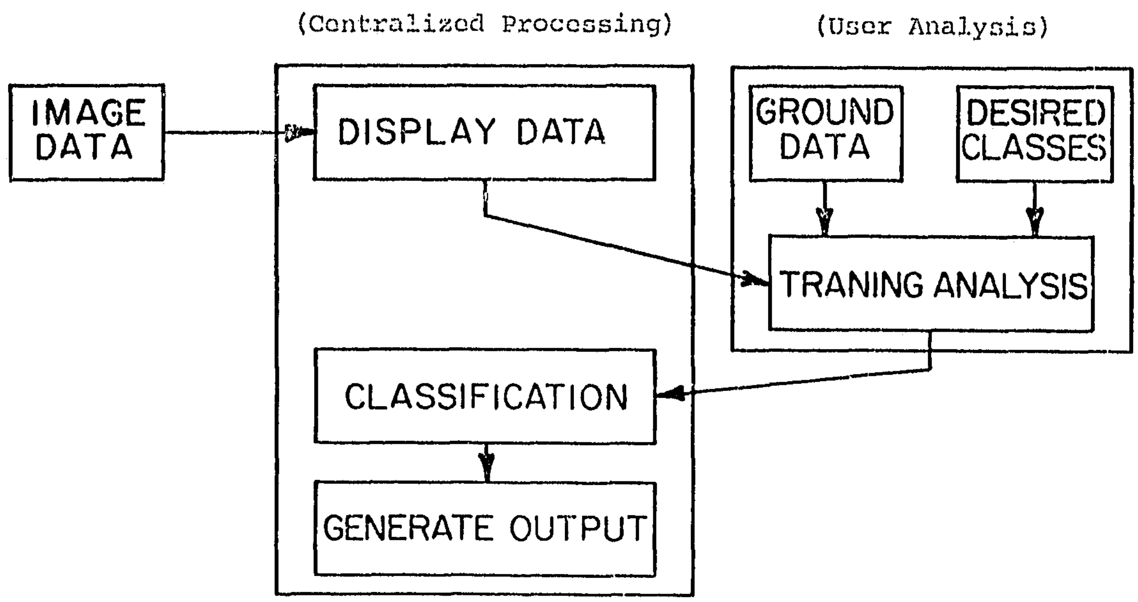
4. Reduce the time pressure on the user frequently caused by the need to tightly schedule his or her time and to schedule the expensive interactive analysis equipment. Taking sufficient time to analyze computer output during the classification process is a critical element too frequently missed in digital analyses. In addition, the distributed approach would give the user the option of suspending the analysis if it is deemed necessary to revisit the field or obtain other reference data.

5. Eliminate the need to invest in specialized computer equipment, and to handle the staffing and maintenance requirements of such equipment.

Thus, our effort is oriented not towards developing new quantitative classification techniques, but rather toward finding ways to implement current techniques in a way better suited to non-research applications.



(a) Centralized Data Analysis Process.



(b) Distributed Data Analysis Process

Figure 1. Comparison Between Centralized and Distributed Analysis Concepts.

monitor, and flexible disk drive. The system can communicate over a high speed phone link to the Universtiy computer system.

In addition to being used as a development tool for the hard-copy based analysis techniques, this equipment will be explored as an additional tool for decentralized data analyses where modest equipment investment is feasible. Microcomputer-based systems like the Visacom have considerable potential for enabling equipment to be more widely distributed, particularly as their cost continues to decline. We will evaluate the pros and cons of this approach as the research continues.

4. Procurement of coordinate digitizer. Again using in-house funds, a digitizing tablet has been obtained which will be connected to the microcomputer system. Software will be developed to enable supervised training site entry, interactive masking of image data to restrict classification to subscene areas, and test site entry for detailed accuracy assessment. The first and last of these activities are currently being explored using an IDIMS system in a cooperative project with the EROS Data Center. This work is described on page B9 report.

In addition to the progress described above, we have been fortunate in the past six months to have been involved in several Federal-State cooperative projects which have had input to this research. Agencies involved include:

1. NASA Eastern Regional Applications Center (ERRSAC). In the past year, ERRSAC began a program of remote sensing technology transfer to state and local governments. Minnesota was one of the first states

selected to be involved, and the Remote Sensing Laboratory has participated in the activities in a technical assistance role. This involvement has included giving a seminar on image processing to state employees, participating in a week-long training course at Goddard, and attending a conference organized by ERRSAC. The ERRSAC training course and conference have provided attitional exposure to the problems involved in training new users. They have further confirmed the potential utility of a decentralized approach to digital classification.

2. <u>Earth Resources Laboratory, NASA Space Technology Laboratory</u>. As an outgrowth of the ERRSAC activity, we have provided consultation to the Land Management Information Center of the Minnesota State Planning Agency. This group has operated a digital geo-based resource information system for the state for ten years. They have recently obtained funds to purchase specialized computer hardware to further support and enhance that activity. Part of that equipment will be for remote sensing image processing, on which we have provided technical advice. As part of that activity, a trip was arranged for a demonstration of the ELAS image analysis system at ERL in Mississippi. This demonstration provided an interesting comparison to the IDIMS system used at Goddard. In particular, the effective use of an unsupervised classification technique was impressive. Applications staff at ERL reported almost exclusive use of this classification technique. We hope to integrate such an option in our decentralized processing approach.

3. <u>Data Analysis Laboratory, EROS Data Center</u>. We are currently participating in a cooperative project with EROS and Minnesota State

Planning. The project is intended to investigate the incorporation of LANDSAT-derived data into the statewide land information system. A wide range of analysis techniques was proposed, and we were able to expand this to include the techniques which we have been exploring under this grant, particularly the use of a coordinate digitizer to enter supervised training set polygons and test areas for accuracy assessment. Subsequent computer work will be performed on the IDIMS image processing system at EROS. This should prove very advantageous to us, since it will permit the various analysis strategies to be tested without the time consuming task of software development. The results of the comparisons will then allow us to concentrate our future efforts on those techniques which look most promising.

The demonstration project is being performed on a study site located north of the city of St. Paul. The area under analysis covers part of four 7.5 minute quadrangles containing a range of urban fringe land uses, from medium density suburban residential to large scale agricultural. Image data for the area have been extracted from an April, 1979, Landsat 3 scene. The system corrections to this image by the new NASA Master Data Processing Facility include a geometric resampling to the Hotine Mercator Projection. The data analysis and processing techniques which are being compared in the study are:

Training techniques:

a. Supervised training using CRT display and joystick to enter polygons

b. Supervised training using color hard copy and a digitized tablet to enter polygons

c. Unsupervised classification

d. Supervised training using digital land cover data currently in the Minnesota Land Management Information System.

Classification techniques

a. Minimum distance to mean

b. Maximum likelihood with threshold

c. Canonical transformation with minimum distance to mean

d. Post-classification smoothing algorithms

Geometric correction techniques

a. Use of NASA corrected data resampled to Hotine Mercator Projection

b. Resampling to convert from Hotine to UTM projection prior to classification

c. Same as (b) but resampled after classification

Data resulting from combinations of the above techniques will be compared to a photo-interpreted reference data set prepared by the Remote Sensing Lab. The reference data consist of six randomly selected photographs covering approximately one quarter of the study area. The interpreted reference data are being geocoded at EROS to enable automatic, full pixel comparisons to the test data sets.

Currently, about half of the above work has been completed.

FUTURE ACTIVITIES

Assuming continued funding under this program, the following areas will be pursued in the coming year:

1. Basic software development on the microcomputer/display system. This will involve programs to transfer data from the University Computer Center, provide optimized disk data storage, and display the data. Down-loading of image processing programs will be attempted. Interface software will be developed to operate the digitizer tablet.
2. Completion of the EROS cooperative project. As has been mentioned, these results will be used to define the direction of our development efforts. At present, it appears likely that this will involve additional attention to unsupervised analysis.
3. Implementation of software based on the results in (2) but designed for decentralized operation.
4. Sample projects to test the decentralized approach using field personnel with limited digital image analysis background. Several prospects for sample projects are available within Minnesota, but attempts will be made to arrange additional demonstrations through NASA ERRSAC and/or EROS.

We feel that the activities of the past six months have put us in a good position to develop the proposed alternative analysis approach. Our involvement with the technology transfer groups at NASA and EROS have further emphasized the needs which formed the basis for our proposal last year. In addition, these contacts should improve our ability to publicize and acquire feedback on the alternative techniques we develop. The exposure to analysis systems at NASA ERRSAC and ERL and the hands-on experience at EROS have brought us up to date on the state of the art. This is critical to avoid "reinventing the wheel." Finally, the availability of the microcomputer and color display system will permit the investigation of a potentially important alternative approach to decentralized analysis.

In short, we feel we are making substantial progress toward defining the technological components of an alternative approach to digital classification which will improve the accuracy and specificity of natural resource inventories made with digital image data. During the next year we will begin testing this system in the "real world".

N81-12508

SECTION C

SYNERGISTIC RELATIONSHIPS AMONG REMOTE-SENSING AND GEOPHYSICAL MEDIA: GEOLOGICAL AND HYDROLOGICAL APPLICATIONS

Dr. Joseph E. Goebel
Dr. Matt Walton
Mr. Lawrence G. Batten
Minnesota Geological Survey

INDEX