123175

QUARTERLY REPORT

Reporting Period: October 1, 1975 - December 31, 1975
Contract Number: NAS9-14016
Title of Investigation:

Research in Remote Sensing of Agriculture, Earth Resources, and Man's Environment

Principal Investigator:

D. A. Landgrebe
Laboratory for Applications of Remote Sensing
Purdue University
West Lafayette, Indiana 47906

Submitted to:

NASA Lyndon B. Johnson Space Center
Applied Physics Branch
Attn: A. E. Potter, Mail Code TF3
Houston, Texas 77058
Mark for: Contract NAS9-14016

2.1 Layered Classifier Adapted to Multitemporal Data Sets

The Multitemporal/Multisegment Problem. Experimental results obtained early in this quarter (Finney County, Kansas) demonstrated that changes with time in spectral response and spectral class definition are so great even within a single segment that there was little to be gained by continuing to restrict our experiments to the multitemporal/single segment case. We therefore decided to proceed with the more general case in which, say, training data is drawn from segment A at time 1, segment B at time 2, and the objective is to classify segment C which is assumed to have data available at both time 1 and time 2.

The procedure we have developed depends on the following assumptions:

1. Segments A, B, and C belong to the same spectral stratum.

2. The informational classes in segment C (the segment to be classified) may be assumed to be a subset of the union of the classes in segments A and B.

3. No dramatic changes occur in conditions during the acquisition of data which would cause the time 1 data for segment C to differ in spectral character from the time 1 data for segment A; and similarly for the time 2 data (segments C and B).

These appear to be reasonable assumptions which are not likely to be relaxed in the future unless they turn out to be unrealistic and a feasible approach is suggested for circumventing them.

Essentially, the procedure is to develop training classes and statistics for each time/segment independently, design the lower levels of the tree (the decision nodes nearest the root) using the training data from one time/segment, and then use the training data from the second time/segment to resolve, wherever possible, uncertainties remaining after application of the data from the first time/segment.

As a simple example, the data for time 1, segment A may yield a decision tree (which may be a multilevel tree, of course) with terminal nodes for Bare Soil, Row Crops, Small Grains, and Other. The data for time 2, segment B might resolve Row Crops into Soybeans and Corn or resolve Small Grains into Rye and Wheat.

Although this sounds rather straight forward, in reality the procedure is rarely so simple to carry out. Examples of questions which typically arise in designing the overall tree include: Which time should be used for the "base" (nearest the root) nodes? What level of certainty should be required before deciding that data from additional times should not be used? The answers to these questions are invariably highly problem-dependent.

Two analyses were performed using CITARS data (LACIE data was at the time not yet available):

Training:

Livingston County, June 10, 1973

Fayette County, July 17, 1973

Classification:

Case 1 Shelby County, June 8, 1973

July 15, 1973

Case 2 Huntington County, June 9, 1973

July 15, 1973

In these cases, the time 2 data were used to resolve uncertainties between soybeans and corn which remained after application of the time 1 data.

The results of these analyses were not particularly exciting, being roughly comparable to those obtained from the local recognition/unitemporal CITARS analyses. About the best we can say is that (1) the results were no worse than the CITARS results, even though some signature extension ("non-local recognition") was involved, and (2) an opportunity was provided for gaining a better appreciation for the complexities involved in interfacing multitemporal/multisegment data with layered decision tree classification. Further experimental work is being pursued with LACIE data for which we are somewhat more optimistic.

Difficulties encountered: One shortcoming of the procedure we have described above is that it allows for no explicit interaction among the data from different times. That is, the training statistics used at any decision node must be drawn from a single time/segment and any interactions between times which might assist in the classification cannot be taken advantage of. This restriction arises from the fact that a simple correspondence between classes in different segments (or even the same segment but at different times) cannot be routinely established. We continue to study this problem (which is related to the more general signature extension problem), but we do not now believe a practical, implementable solution exists.

A technical difficulty we are having involves the manipulation of statistics and tree descriptions for the component times/segments, which is proving to be a burdensome and error prone bookkeeping task. It will be necessary to engage in a short programming effort to get this problem under control.

Layered Classifier Software Improvement. The development of software which will enable the use of the layered classifier in a relatively routine fashion comparable to LARSYS usage is proceeding somewhat behind schedule. The original milestone of December 31, 1975 for availability of the basic classifier software has proved to be optimistic, possibly by about a month. The delay is associated at least in part with providing internal (to the classification results themselves) documentation of the decision logic used to produce the results.

Optimal Classifier Design Procedures. A post-doctoral research associate joined this project in early November and has been reviewing previous research as well as becoming familiar with the remote sensing problem in general. He is assigned full-time to this project and we look forward to his eventual contributions.

2.2 Development of Spectral Strata from Clustering Techniques

Spectral stratification activities during the second quarter included work in three areas: (1) development and testing of multivariate pattern recognition (clustering) techniques to determine and delineate spectral strata in LANDSAT MSS data, (2) compilation of physical factor maps and data to use in correlating spectral strata and physical factors and (3) development of procedures to digitize physical factor data and register it to LANDSAT data. A major redirection of the project was initiated in late November at the request of JSC to provide for (1) increased coordination of research by LARS, ERIM, and UCB on the signature extension problem and (2) more timely delivery of results in order to meet LACIE deadlines in the spring of 1976.

Development and Testing of Clustering Techniques. The first four clustering procedures were fully implemented; procedures 5 and 6 are implemented, but presently are limited to 60 classes. Strata maps have been produced for several frames using the procedures. A capability for the procedures to be applied to segments only (rather than entire frames) has been implemented. In addition, the procedures are being modified to accommodate smaller block sizes since the 200 lines x 200 columns block size appears to be larger than some strata.

Strata produced by the different procedures are being evaluated by comparisons of test field recognition resulting from use of local and non-local training segments. An example of the classification results for scene 1689-16382, acquired June 12, 1974 over the central crop reporting district of Kansas is shown in the following table:

Source of Training Statistics*	Segment Classified*	Classification Performance (%)			
		Wheat	Other	Overall	Average
Stafford (1)	Stafford (1)	93.4	86.6	89.9	90.0
	Barton (2)	51.8	69.0	55.4	60.4
	Ellsworth (2)	59.9		59.9	59.9
	Rice (2)	22.0	65.8	50.0	43.9
	Ellis (1)	61.5	64.5	63.4	63.0
Barton (2)	Barton (2)	98.4	84.7	96.0	91.6
	Ellsworth (2)	90.0		90.0	90.0
	Rice (2)	97.4	61.2	74.3	79.3
Rice (2)	Rice (2)	65.9	86.7	79.9	76.3
	Stafford (1)	66.8	76.3	70.7	71.5
	Barton (2)	43.3	97.6	54.9	70.4
	Ellsworth (2)	31.2		31.2	31.2
	Ellis (1)	50.4	48.9	49.4	49.7
Ellis (1)	Ellis (1)	11.5	84.7	58.5	48.1
	Stafford (1)	60.0	54.3	57.6	57.1
	Barton (2)	50.1	66.9	53.7	58.5
	Ellsworth (2)	62.7		62.7	62.7
	Rice (2)	19.5	53.2	41.0	36.4

* Strata number; segments with the same number are in the same strata.

The results appear favorable for the segments within strata-2, i.e., similar classification performance was achieved for local and non-local training within the strata. Signature extension however was not successful for the segments within strata-1.

These tests have been severely limited by both the quantity and quality of data available. Some of the problems include: only one segment per stratum, too few fields per segment, some segments have only wheat, and mis-registration which limits multitemporal usage. As a result of these data set limitations we have prepared a recommendation for the acquisition of a new data set in 1976 in which these characteristics would be eliminated or at least minimized.

By the end of December we had transmitted a possible partioning of the segments in the southwest crop reporting district of Kansas to ERIM and UCB.

<u>Physical Factor Data</u>. The second area of activity has been to complete the definition of the physical factors which will be correlated with the spectral strata found in the LANDSAT data. The factors are: soil associations, soil color, crop growth stage, yield, precipitation, temperature, and agricultural land use. Data and maps have been acquired for most of the factors for Kansas for 1973-74. A major technical question is how many classes each of the primary factors should be divided into to achieve the most meaningful relation with the strata found in the LANDSAT data. We have concluded that this cannot be readily determined a priori and are therefore working on assembling the data so that several classes and combinations of classes can be tested for each factor. However, work on correlating physical factor data and LANDSAT strata has been delayed by the redirection and increased emphasis on the development of partioning techniques.

<u>Ancillary Data Registration</u>. The task defined for this portion of the spectral strata study is to digitize and register several physical variables from maps and other sources with LANDSAT data for appropriate LACIE sites. This task is being conducted in conjunction with the ancillary data registration research task with theoretical considerations being made in the research task and practical applications being made in the spectral strata task.

Consideration of the expected relative importance of various physical factors was made and the following were defined: soil association, soil color, crop calendar, temperature, precipitation, and agricultural land use. These factors were broken into a number of levels and maps and tabular data were compiled for the test site areas.

The process of transforming the map and tabular data into a digital tape file registered with LANDSAT data consists of several steps:

1. Compilation of map to be digitized from map or tabular data.
2. Digitization of map contours and polygons plus control points using a coordinate digitizing table.

3. Transformation of contour, arc and point data to a uniform grid of data values.

4. Registration of the gridded data to a reference LANDSAT data set.

5. Optional geometric correction and transformation of the aggregate data set to a map projection.

A generalized soil map for the state of Kansas has been digitized on a table digitizer and the coordinates of arcs and nodes punched on cards. This activity was completed by late October and work began on algorithms to transform the data. Software already in existence developed by other projects was adapted to the present problem. In the case of polygon type data such as that generated from the soil map a relationship must be made between each arc and the categories separated by the arc. The areas enclosed by the arcs are then filled with numbers identifying the category. A second channel is generated with the arcs only (boundaries) delineated for use by interpreters.

The digitized control points from the map form an additional input which is used to define a warp function to overlay the map data onto a LANDSAT frame. Work has proceeded to the point of getting control points specified and completion of the overlay for the soil map is expected by early January 1976. Maps of the remaining physical factor variables will be digitized and registered during the next quarter.

2.3 Field Measurements for Remote Sensing of Wheat

Field measurements activities during this quarter have included data acquisition, data processing and evaluation, management of the data library, data analysis, initiation of experiments for analogous vegetative areas in the U.S. and Soviet Union, and project leadership and coordination.

Data Acquisition. The primary data acquisition activity during this quarter was the establishment of a capability to acquire, process, and duplicate aerial photography on LANDSAT-2 overpass dates for which NASA would not be flying data acquisition missions, i.e., weekends and holidays. The photography will be the same scale, type, and format as currently obtained during the NASA C-130 flight missions. Photography of excellent quality was obtained on November 8, 1975 over the Finney County test site in a test of the system. A proposal to cover the cost of this activity for next year has been prepared and submitted to NASA/JSC.

Data Processing and Evaluation. During this quarter data has continued to be processed as it became available from the data acquisition centers. All of the LARS' field spectroradiometer data have been digitized and reformatted except for the last mission. All of the FSS (S-191) data which has been received from JSC has been reformatted. Receipt of the remaining data has been delayed while corrections were made in the data due to a problem with the dichroic of the thermal detector. Data from the 24-channel MSS system for four missions have been received and processed. Serious problems including bit errors, saturation, and banding have been found in the data.

Data from the interferometer (VISS) have recently been received for the 1974 missions; however, JSC subsequently reported finding an error in it which will require preparation of new data tapes. No data, either spectral or agronomic, have been received for the missions performed by ERL. ERL has had to perform an extensive analysis of the spectral data to correct it for several instrument malfunctions; however, requests for the ground truth data so that we can get it punched on cards have not been filled.

Data Library. Procedures for cataloging, storage, retrieval, and distribution of the field measurements data have been established. Field measurements data of all types have been received by the data library. All available data, including inventory and periodic observations, meteorological and optical depth data, data logs, photo data for LARS missions in North Dakota, and computer tapes with Exotech Model 20C and FSS (S-191) data, have been distributed to researchers at Texas A & M University and ERIM. Canopy modeling data have been sent to Colorado State University and ERIM.

Data Analysis. Analysis activities have included work on the thermal data acquired by the FSS over the intensive test site in North Dakota and data acquired by LARS at the experiment station near Williston. The latter data are providing an indication of the utility of the spectral measurements, the angular distribution of emitted radiation, and temperature profiles of wheat canopies. The analysis is providing information which is being used for planning further investigations.

Data from an experimental procedure utilizing a laser probe technique have been reduced and the results used to study a wheat canopy. The technique shows considerable promise as a means of characterizing the physical characteristics of crop canopies which when combined with incident radiation measurements will provide valuable information about the interaction of solar energy and crop canopies.

A small offset problem between the computed reflectances for the silicon and lead sulfide channels of the Exotech Model 20C has been noted. The instrument and the data have been analyzed in an effort to determine its cause and a possible data correction for it.

Remote Sensing Experiments for Analogous Vegetative Areas in the United States and Soviet Union. In October we were asked to participate in the joint U.S./U.S.S.R. experiments on remote sensing of vegetation and perform the same functions as in the field measurements project. Our involvement to date has included selection of flightlines for the Hand County, South Dakota test site, a visit to the test site to assist in data acquisition, preparation of the data acquisition and exchange plan, description of the format of data to be used for exchange of data, and preparation of ancillary data for exchange in January. We are awaiting arrival of the early November mission of 24-channel MSS data, FSS data, multispectral photography, and ground observation data. This data will be prepared for exchange in March. We are also beginning preparations for the spring and summer data acquisition missions; this activity has included selection of test fields, definition of data acquisition procedures, and designation of mission dates.

Project Leadership and Coordination. Primary activities this quarter have been continued communication with personnel at JSC and other centers concerning the project. Problems which have received particular attention are scheduling of 1976 missions, data processing problems and schedules, operation of the VISS for 1976, and data quality evaluations.

2.4 Scanner System Parameter Study

Major Activities. During the first part of the Second Quarter it was decided to use the LACIE supersites as the data sources for the spectral data ensemble which will serve as the data sources for this project. The S-191 fast scan spectroradiometer was chosen as the principal data source for this study. As a part of the field measurements research project (Task 2.3) the S-191 data from the LACIE supersites were routinely being reduced for analysis. A segment of data from the Kansas supersite was selected for use to develop a trial spectral data ensemble. Appropriate software was written and a spectral data ensemble was prepared. This ensemble is intended to be used in checking the functional expansion coefficient characteristics software that is being written.

In addition, a number of tasks directed to the solution of the problem of digitally representing a multispectral scanner system for design evaluation have been identified. They are: 1) A classifier spectral error model; 2) A scene spacial characteristics model; and 3) A scanner physical parameter simulation.

Work has principally focused on the classifier spectral error model. A literature survey has been conducted on multiclass classification error models to set the ground work for developing a suitable approach for this study. No model apparently exists for computing the classification error on a given set of classes from their statistics. Various approximation schemes are being surveyed for use in the scanner study and preliminary considerations have been made on the spatial scene model and scanner model but no results have yet been produced.

At the request of JSC, LARS has agreed to assist in the development of additional quantitative support for the selection of engineering specifications of the thematic mapper instrument that is to be a part of the LANDSAT-D satellite. Such a study is related to some of the basic elements of the scanner system parameter selection project and, therefore, the thematic mapper study was placed in the scanner system project at LARS. The primary data source for the study is the MSDS multispectral scanner system operated by NASA. Data from flightlines over the Finney County, Kansas and Williams County, North Dakota LACIE supersites were selected by NASA for preprocessing and analysis. The MSDS data will be spatially degraded to simulate satellite data and spectral bands will be chosen that most nearly represent those chosen for the thematic mapper.

Special software for handling the MSDS data and the spatial degregation routines has been developed and tested and preliminary nalyses of the data are proceeding. GISS, ERL, and NASA/Goddard are participating in the analysis of the data with LARS and LARS will be furnishing appropriately preprocessed data to these laboratories for analysis. It will be necessary to set back some of the milestone goals for the originally proposed scanner system study in order to accommodate this special analysis project. However, many of the results of this project will be usable in the basic scanner system parameter study.

Technical Problems and Solutions. The MSDS data that are being used in the special thematic mapper simulation study have exhibited noise, bit error and saturation problems. Work is under way to reduce the impact of these problems but their effects will not be known until some preliminary analytical results have been evaluated.

Data Analysis. Data from both the field spectral system and the S-191 airborne spectroradiometer have been preprocessed and used to prepare the spectral data ensemble mentioned above. MSDS tapes have been reformatted, sun angle corrected, and are being prepared for analysis.

Plans for the Next Quarter. During the next quarter no further work will take place on handling the spectral data ensemble as all activities will be concentrated on the production and analysis of the TM simulated data. Several flightlines will be spatially degraded and classified.

2.5 Transfer of Computer Image Analysis Techniques (Remote Terminal)

Activities under this section of the work statement include continued support of the JSC remote terminal and the development of technology transfer activities and materials. During the first quarter, milestones for the year's activities were identified and reported. Three major milestones were identified for this quarter: Materials Development, Interim Training Programs and Remote Terminal Support. Each of these activities are described in detail below.

Materials Development. The main thrust was in the design of a simulation exercise dealing with the analysis of LANDSAT data using computer-aided analysis techniques. Designed to enable the reader to gain an appreciation of the decisions and trade-offs made by an experienced analyst, the publication describes the sequential process of analyzing a LANDSAT data set and emphasizes the interaction between man (analyst) and machine (computer). Typical products (results) of each step in the analysis are shown. The analysis objectives of the simulation exercise deal with Forestry so that the document will serve as a supplement to the LARS Forestry Applications Project task as well as a model for future simulation exercises dealing with other applications.

The current status of this effort is that a draft, including photographs and copies of computer-generated output, has been completed and the materials are ready for internal evaluation. At least six people will go through the simulation exercise and provide feedback to the authors. After necessary revisions to improve the educational quality the document will be published and distributed.

Other educational materials, such as a LACIE Simulation and a Regional Applications Project Simulation, are being considered and planned. Additional materials development will be dependent upon the need for and scope of Interim Training Programs (described in the next section).

Interim Training Programs. The last quarterly report recommended a training program for 8 to 10 JSC personnel at LARS. These personnel would serve as instructor-consultants during additional training programs at JSC. Their services would be invaluable in conjunction with LARS personnel if, for instance, a need arose to train additional LACIE analysts at JSC, or for other on-site training needs which might arise in the future. Although no request for this training activity has been received, preliminary plans have been made.

Remote Terminal Support. Support of the JSC remote terminal was continued during the second quarter of the contract period.

Summary. No problem areas arose during the second quarter. Activities under this task have progressed according to the milestone schedule.

2.6 Research in Remote Sensing Technology

The non-LACIE Research task has three activities as defined in the goal and milestone plan. These are: 1) Research on Ancillary Data Registration Techniques, 2) LANDSAT LACIE Site Imagery Enhancement, 3) Soil Reflectance Research. The techniques under study in (1) although of a general nature are being applied directly to the Spectral Strata task (2.2) to register ancillary data relevant to wheat signature extension. Task (2) is a continuation of a LACIE Image Enhancement task from CY75 and has as its goal the development of software for enhancement based on algorithms developed in CY75 and delivery of example enhanced imagery. An enhanced imagery evaluation subtask was also included to complete work started in CY75. The third task is a soil spectral response study which is using the EXOTECH field spectrometer and the goal is to relate detailed soil spectral response to LANDSAT spectral response for the same material. An additional task has been included, at JSC request, to register LANDSAT data from LACIE sites for evaluation of registration error effects.

<u>Ancillary Data Registration</u>. One important motivator for ancillary data registration derives from the signature extension problem. Physical factors are assumed to cause variation in spectral response which invalidates training parameters at observation points distant from the training area. Digital registration of physical factor data with the remote sensing data will enable quantitative evaluation of their effect. The same reasoning applies to many other classification problems and for these reasons an ancillary data registration task is being pursued.

The types of physical data which are most often encountered are in the form of maps. The data is usually originally obtained in tabular or photographic form and compiled into a map. The formats of the data on the map includes contours, lines, polygons and points. Figure 1 graphically describes some of these forms. It is assumed that the physical data is obtained in a map format and the goal of this study is to define procedures for transforming the map into a digital image-like record registered pixelwise with remote sensing data. The basic steps in the map transformation process are assumed to be:

1. Digitization of the physical map.

2. Conversion of the lines and points to a uniform grid of points.

3. Registration of the gridded data to a reference data set.

Each step involves a number of parameters and algorithms which comprises the elements of the study. Some of the considerations are:

1. Density of samples for contours and arcs.

2. Effects of position errors.

3. Scheme for defining interiors of polygons.

4. General considerations of the information content (bandwidth) of a map.

5. Interpolation algorithms for gridding contour and point data.

6. Coordinate systems and warping functions for registration.

These and other factors are being considered in the context of the variables of interest for spectral strata definition. The variables currently under study are:

1. Soil Type and Association.

2. Temperature and Precipitation.

3. Slope and Elevation.

4. Cropping Practice.

Current status of the project is that a soil map for the State of Kansas has been digitized on a table digitizer using a particular set of parameters. Polygon gridding software has been adapted for use in this study and is being checked out. A CALCOMP plotter routine has been written and the digitized soil map has been reproduced for checkout purposes. Some problems have been encountered and the plotter routine is being debugged. The next step will be to identify control points for a full frame of LANDSAT data corresponding to ones defined on the map. A registration polynomial will then be defined and overlay attempted. The first map transformation is using known and available procedures and algorithms as a test. The results will be evaluated and further plans and goals set on the basis of the results. Registration of the soil variable is expected by mid-January 1976. The next physical variable to be considered will be precipitation.

LANDSAT Image Enhancement. The image enhancement task was to have been completed by October 1 by the delivery of an algorithm for performing the instantaneous field-of-view (IFOV) compensation filtering developed at LARS during CY75. Problems in programming an efficient multichannel version of the convolution algorithm delayed completion of the program. A program is now operational and performs a four channel enhancement nominally five times faster (in terms of job time) than the previous version. The previous version did only one channel and required extensive tape and disk buffering. The new convolution algorithm uses a multichannel cylindrical buffer and computes all points in one output data line in one pass rather than two.

The basic process of convolution as implemented in the current IFOV algorithm is to map each point of the input image to a four by three block of points in the output image. The points in the output are a weighted sum of the original point and four points to either side and above and below it.

If we let $P_{0,0}$ be the original point it will generate a $\hat{P}$ array as follows:

$$(P_{0,0}) \qquad \begin{bmatrix} \hat{P}_{0,0} & \hat{P}_{0,1} & \hat{P}_{0,2} \\ \hat{P}_{1,0} & \hat{P}_{1,1} & \hat{P}_{1,2} \\ \hat{P}_{2,0} & \hat{P}_{2,1} & \hat{P}_{2,2} \\ \hat{P}_{3,0} & \hat{P}_{3,1} & \hat{P}_{3,2} \end{bmatrix}$$

$$\hat{P}_{n,\ell} = \sum_{j=-4}^{4} [C^{(v)}_{n,j} \sum_{k=-4}^{4} C^{(h)}_{1,k} P_{j,k}].$$

The two C matrices $C^{(v)}$ and $C^{(h)}$, represent the weights of input points in calculating the new point. $C^{(h)}_{1,k}$ represents the weight of $P_{j,k}$ in calculating an intermediate point $\tilde{P}_{j,1}$. $C^{(v)}_{1,k}$ represents the weight of $\tilde{P}_{k,j}$ in $\hat{P}_{1,j}$.

The two dimensional nature of convolution facilitates its partitioning into horizontal and vertical components. This implementation reads a line of data and horizontally convolves it to form a $\tilde{P}$ line. This horizontal process is repeated until sufficient P lines have accumulated to allow the calculation of a set of $\hat{P}$ lines. The $\tilde{P}$ lines are kept in a cylindrical buffer. After each P calculation, the oldest line in the buffer is discarded and a new $\tilde{P}$ line is read in. The buffer, therefore, moves like a rolling cylinder through the input image.

Each $\tilde{P}$ can be calculated by the following

$$\tilde{P}_{1,k} = \sum_{j=-4}^{4} C^{(h)}_{k,j} P_{1,j}$$

Each $\hat{P}$ can, in turn be calculated by

$$\hat{P}_{1,k} = \sum_{j=-4}^{4} C^{(v)}_{1,j} \tilde{P}_{j,k}$$

In the flow chart in Figure 2 INLINE is the P line, CYLINDER represents the $\tilde{P}$ lines and OUTLINE represents the $\hat{P}$ lines. NSAMPLES is the number of points per P line.

Directly computing $\hat{P}$ and P would require the following time (ignoring stores and loads):

$$T_D = 9 \cdot (t_a + t_m) \cdot 9 \cdot N_{lines}$$

$$n_{columns} \cdot 12$$

$$= 982(t_a + t_m)\ n_{lines} \cdot n_{col}$$

where t_a is CPU time for addition,

t_m is CPU time for multiplication,

n_{lines} is number of input lines,

and $n_{columns}$ is number of input columns.

Partitioning the process into components gives the following time:

$$T_p = 3 \cdot n_{lines} \cdot n_{cols} \cdot 9(t_a + t_m) +$$
$$4 \cdot (3 \cdot n_{lines} \cdot n_{cols}) \cdot 9(t_a + t_m)$$

where the terms represent horizontal and vertical times respectively.

$$T_p = 135(t_a + t_m)\ n_{lines} \cdot n_{col}$$

Computing the percentage improvement gives the following conclusion:

$$I = 100\% - \frac{T_p}{T_D} \cdot 100\% = 100\% - 13.7\%$$
$$= 86.3\%$$

In summary, the convolution algorithm generates for each input point a 4 by 3 block of output points. Each output point is a weighted sum of some group of input points. By calculating an intermediate group of horizontally computed points a net savings of 86.3% over directly computing each output point can be recognized in computer time.

The algorithm is being run on test imagery from Montana and Kansas. A problem in specification of the coefficients is suspected in the new algorithm and this is under investigation. Also, a cubic interpolation function is being added so that IFOV and cubic enhanced images can be compared. Discussion with our JSC contract monitor Dr. Potter resulted in the recommendation to continue refinement and production of test output at LARS until the value of the enhancement is shown rather than to transfer the software to JSC as indicated in the milestones. Enhanced test data of LACIE sites will be delivered to JSC as it becomes available on tape and in polaroid image form.

Temporal Registration of LACIE Sites. An image registration task was requested by JSC which is currently in progress. Eight tapes containing LANDSAT data for four dates for two sites were received in October. Severe overload on the image registration system and personnel has caused delays in processing. The registrations are currently in progress and are expected to be completed by early 1976.

Spectral Reflectance of Soils. The research outlined under this study is proceeding at a satisfactory rate toward the fulfillment of its objectives. The results, although inconclusive at this point, are very promising.

Efforts to reveal the nature of the absorption band appearing in the Exotech spectral data at 1.4, 1.9, and 2.2μm have at this time resulted in the intensity and spacing being attributed to inter layer moisture in the clay fractions of the samples. More study is forthcoming as an examination of the validity of these findings.

Satisfactory progress is being made in the search for specific regions wherein the various soil parameters (Cation Exchange Capacity (CEC), Organic Matter Content, Silt Content, Free Iron Oxides, etc.) explain a predominance of the observed spectral variation. As in the previous case the present results lack confirmation and require additional testing.

Progress in evaluating the role of organic constituents in imparting spectral character to soils has been impeded by the inaccessability to the results of organic constituent analysis being performed on the samples sent to another lab.

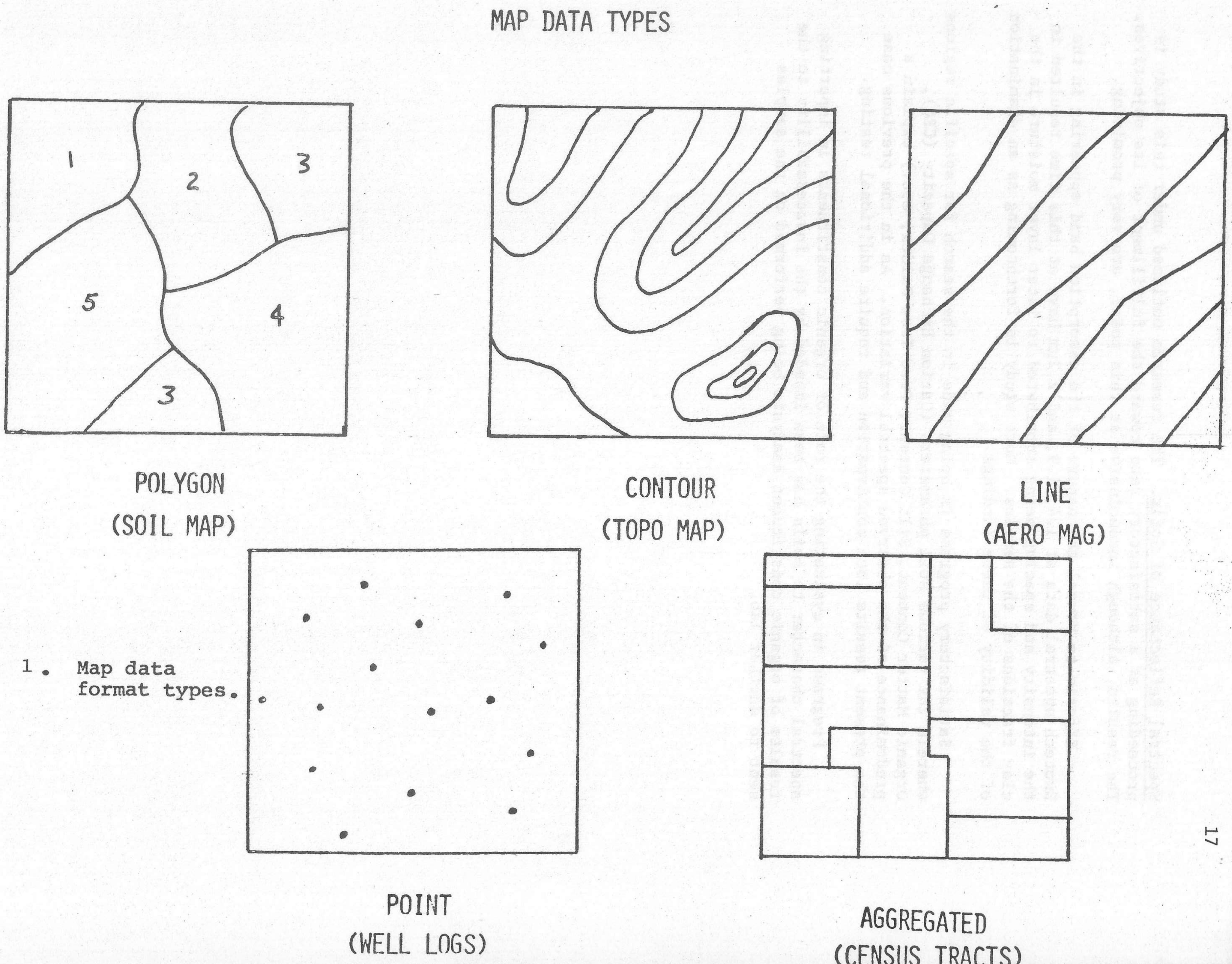


Figure 1. Map data format types.

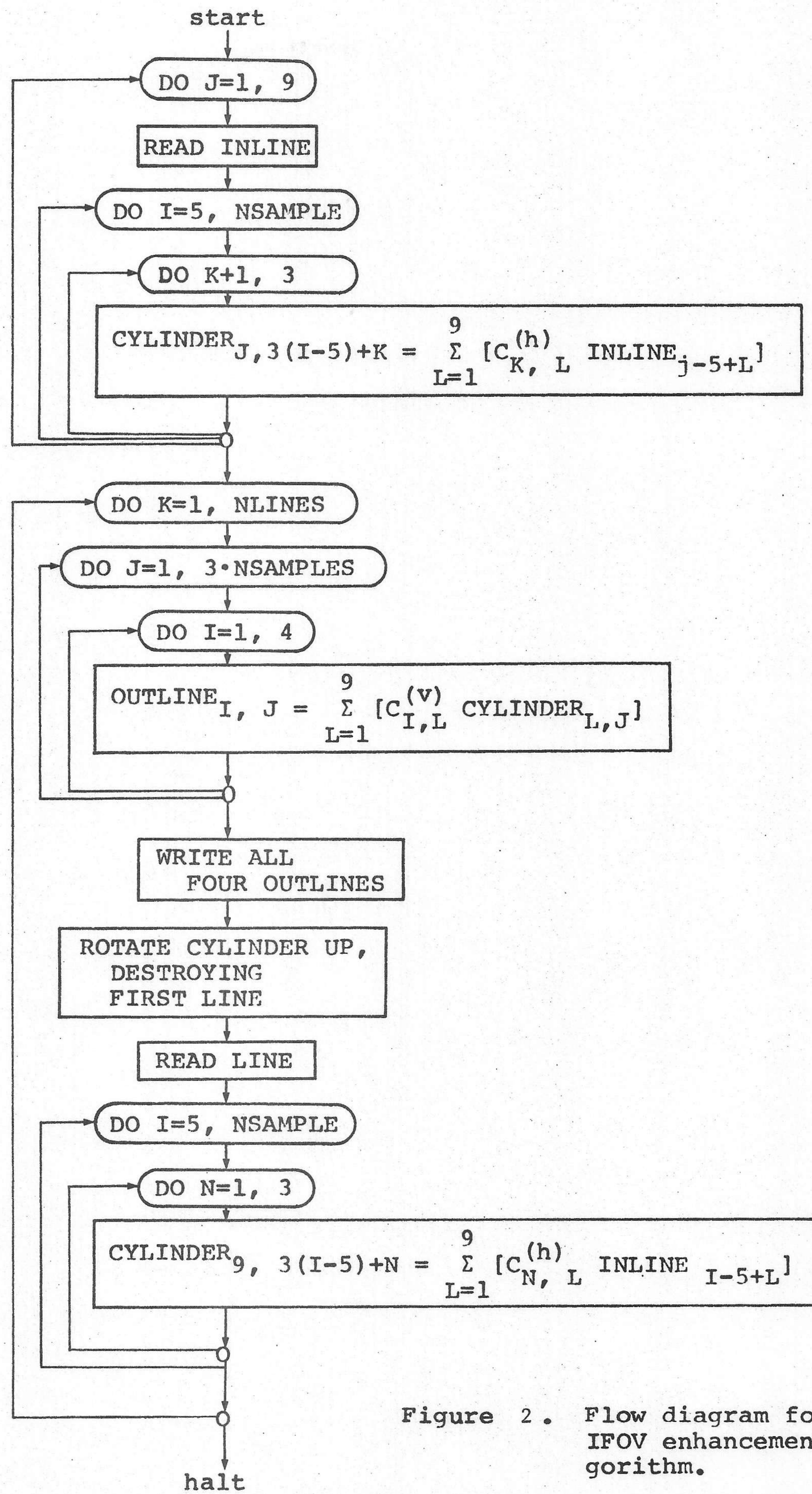


Figure 2. Flow diagram for IFOV enhancement algorithm.

2.7 Forestry Applications of Computer-Aided Analysis Techniques

Introduction. Work at LARS is supporting the FAP by documenting and refining computer-aided analysis techniques for forestry applications. Under this broad objective, specific subtasks of 1) defining training procedures, 2) statistically evaluating results, 3) selecting optimum wavelength combinations, 4) analyzing multitemporal data, and 5) evaluating change detection methods, have been defined. Table 1, a Research Task Matrix, gives a general indication of the status of each task.

In addition, the opportunity to improve the classification map output was identified and is being investigated.

Accomplishments to Date. The major effort during this period has been in producing documentation for the first task, i.e., training area selection. A draft document has been prepared which gives an in-depth description of the procedures defined as the "modified" or "interactive" cluster approach. This document is presently being prepared in draft form and will be subjected to internal review at LARS prior to publication.

The activities within the statistical evaluation task have been intensified. Currently, data from the Sam Houston and Hoosier site are being evaluated with the Stage II techniques previously described. Figure 1 is a flow chart which graphically describes the procedures and inputs to the Stage II evaluation techniques. The procedures document has been evaluated, and is being modified to better account for population size in the statistical evaluation.

Work has been completed on a study of the optimum wavelength bands or regions for forestry classifications. This work has been conducted over the San Juan Mountain site with SKYLAB - collected MSS data. This data set was utilized because it was the best data available to the project to meet the defined task. A report outlining this study's results will be available in the near future.

Multitemporal data includes wavelength region and time interaction problems that must be understood prior to analysis. The 7-date LANDSAT data set for the Hoosier National Forest site presents an ideal test for multitemporal analysis. The multitemporal Sam Houston data set is not considered optimum for this study because it is of inferior data quality and narrow time span.

The questions currently being asked concerning multitemporal analysis revolve about the ability of existing separability processors (algorithms) to identify the "best" channels for analysis. Also, it is questionable whether temporal spectral signatures are meaningful since temporal/spectral interactions result in many, complex classes. It may be that a sort routine (e.g., layered classifier) which only uses one date at each sort is more effective. Methods to optimize the best channel selection process are being considered for comparison.

Table 1.

RESEARCH TASK MATRIX

SR&T Research FAP

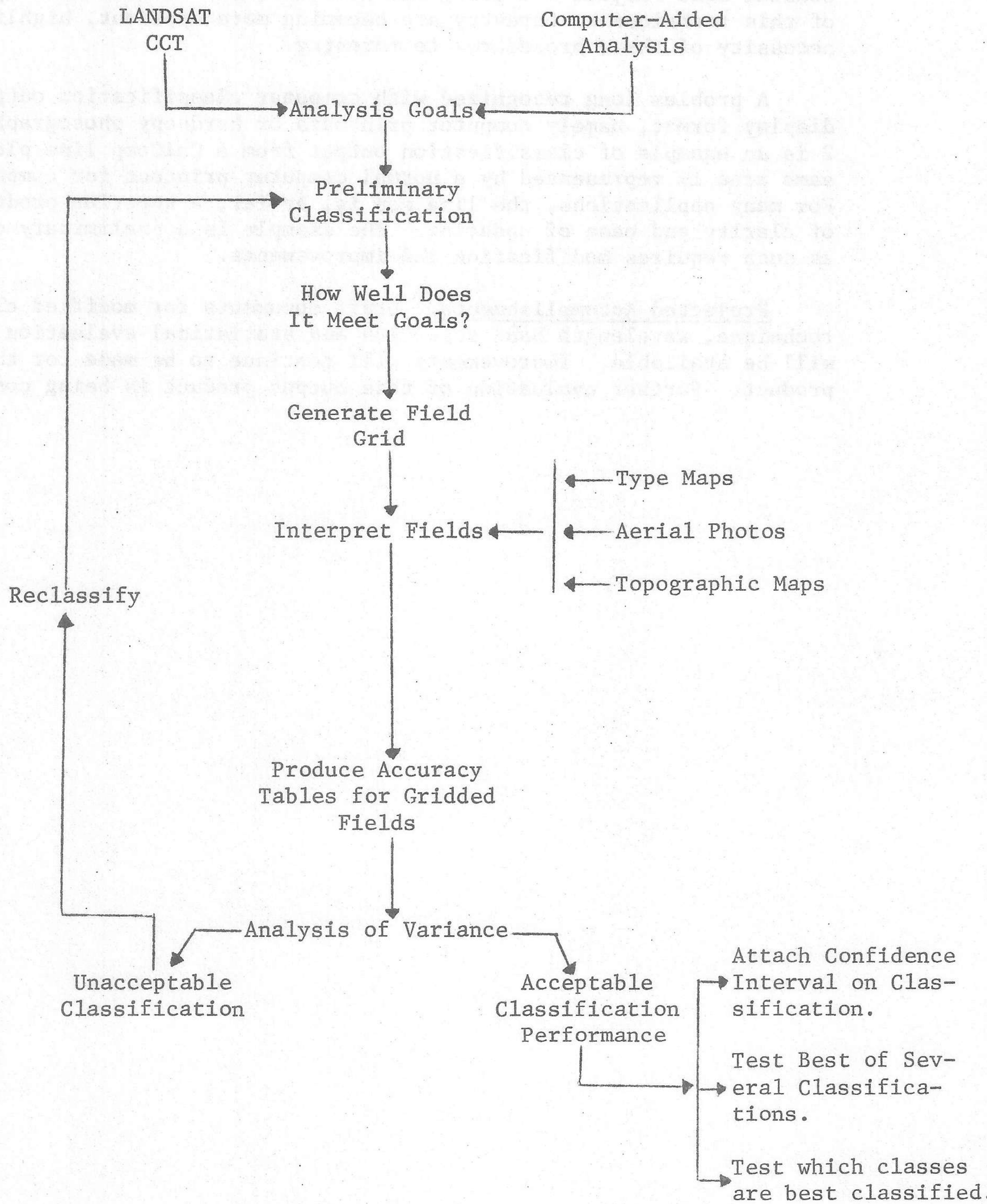
TASKS	Current Research Status: Objectives	Approach	R & D	T & E	Documentation	
1. Training Area Selection "Modified Cluster	X	X	X	X	X	To develop and document procedures for selecting from MSS data training areas for performing forest cover classifications
2. Statistical Evaluation	X	X	X			To develop and document a suggested set of procedures to follow for evaluating computer classifications of forest sites.
3. Wavelength Band Selection	X	X	X	X	X	To select and evaluate which wavelength bands or regions are most suitable for classfying forest cover with MSS data.
4. Multi-Date Analysis Procedures	X	X				Develop a set of procedures which are best adapted for using multitemporal data for classification of MSS data over forested test sites.
5. Change Detection	X	X				Investigate the various approaches to performing change detection with LANDSAT data.
6. Line maps	X	X	X			Develop techniques to produce line rather than grayscale maps.

Multitemporal analysis is an important and even critical part of any change detection work being contemplated as part of this task. Currently, developments in change detection techniques being evaluated by the Texas Coastal Zone Project are being watched closely. Identifiable applications of this technique to forestry are becoming more apparent, highlighting the necessity of these procedures to forestry.

A problem long recognized with computer classification output is the display format, namely computer printouts or hardcopy photographs. Figure 2 is an example of classification output from a CalComp line plotter. The same area is represented by a normal computer printout for comparison. For many applications, the line map is, by far, a superior product in terms of clarity and ease of updating. The example is a preliminary draft and as such requires modification and improvements.

Projected Accomplishments. Draft documents for modified cluster technique, wavelength band selection and statistical evaluation of results, will be available. Improvements will continue to be made for the CalComp product. Further evaluation of this output product is being contemplated.

Figure 1. Steps in Implementing the Stage II Statistical Evaluation Technique

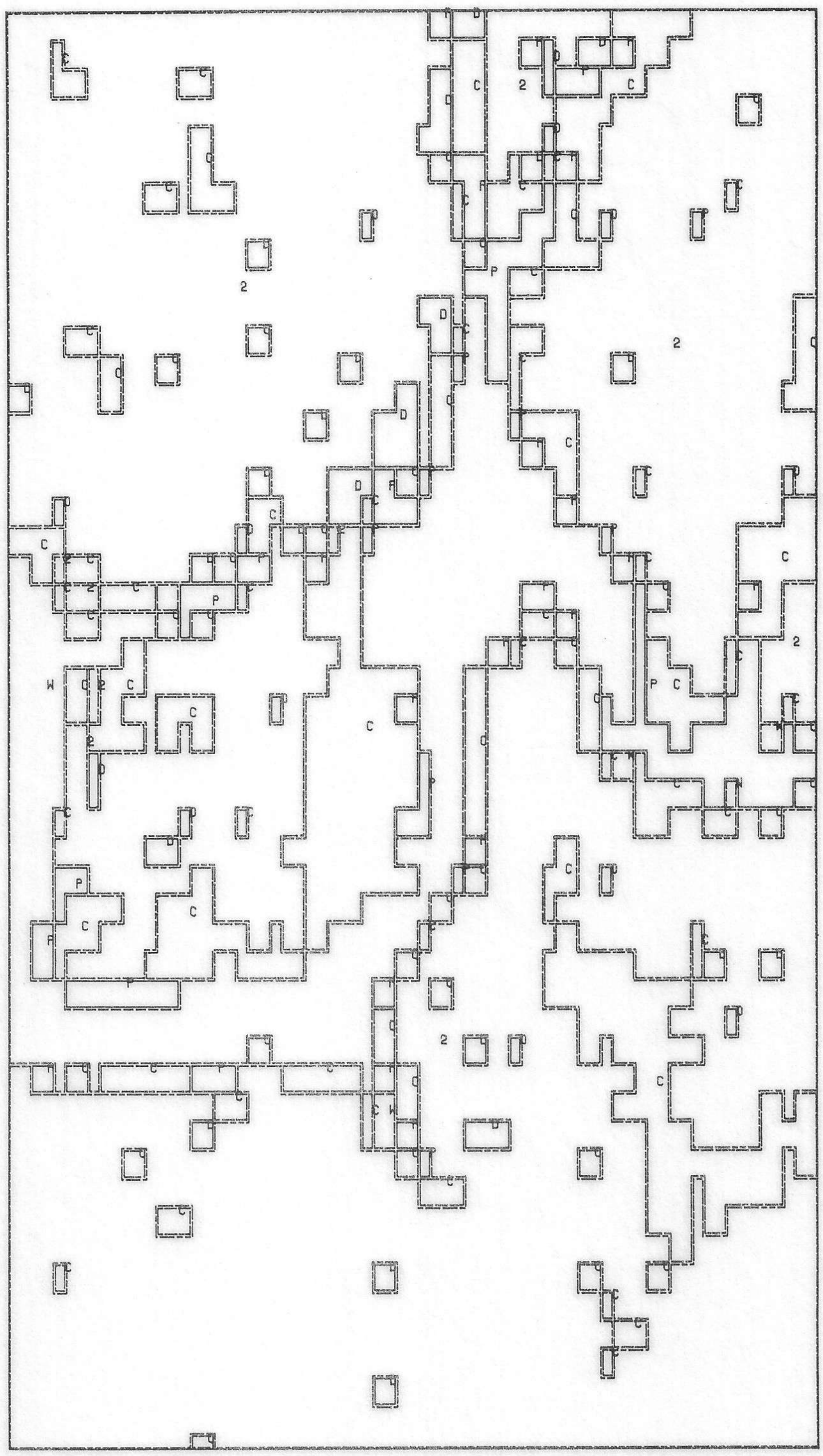


LABORATORY FOR APPLICATIONS OF REMOTE SENSING
PURDUE UNIVERSITY

CALCOMP CLASSIFICATION MAP

CLASSIFICATION STUDY 515658670	CLASSIFIED
RUN NUMBER............ 73051708	DATE DATA TAKEN... 5 / 4 / 73
FLIGHTLINE.... 128516001 IN	TIME DATA TAKEN..... 1000 HOURS
DATA TAPE/FILE NUMBER.. 1970/ 1	PLATFORM ALTITUDE. 3062000 FEET
REFORMATTING DATE. NOV 15,1974	GROUND HEADING..... 180 DEGREES

CLASSIFICATION TAPE/FILE NUMBER ... 537/ 2



P PASTURE
D DECIDE
2 DECIDE 2
W WATER
C CONIFER

CLASSES

SYMBOL	CLASS	SYMBOL	CLASS
C	CONIFER	P	PASTURE
P	PAST-CR-	D	DECID1
W	WATER	C	CR-BR-4
P	CR+RD 3	2	DECID2
P	CR-PAS-3	W	H2ODEEP

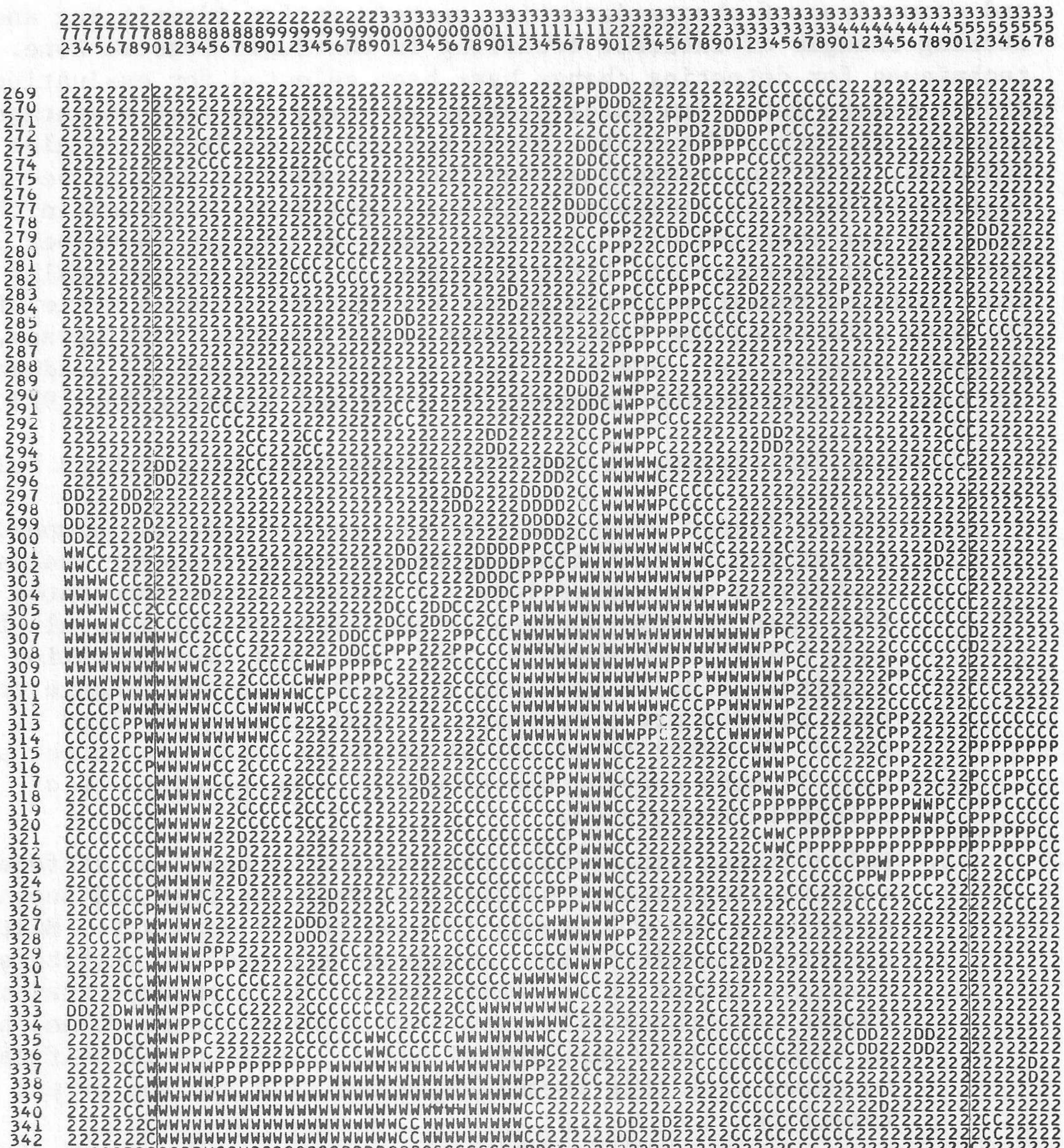


Figure 2

2.8 Regional Applications Project (RAP)

The main thrust of the RAP project for CY76 is the implementation and comparison of change detection techniques for identifying and monitoring changes of interest occurring in the Texas Coastal Zone. Four techniques for detecting change have been selected for evaluation. Direct Change Detection holds promise of being the most advantageous; however, the other three have potential and, therefore, are also being explored. All four techniques require the temporal data to be spatially registered prior to analysis. The data sets being utilized in this analysis are composed of a four date temporal overlay (November 27, 1972; May 8, 1973; December 15, 1973; February 25, 1975) geometrically corrected and precision registered with ground control points at a scale of 1:24,000. The data sets cover the Port O'Connor, Pass Cavallo, Port Lavaca, and Austwell quadrangles. Due to a lack of data the Austwell quadrangle overlay does not include the November 27, 1972 and December 15, 1973 data.

A brief description of the four change detection methods follows:

1. Direct Change Detection - Remote sensor data is independently classified at both times t_1 and t_2 using whatever analysis techniques are appropriate. A direct change comparator is then used to code changes in the classification result for each pixel. For example, if a two class classification is obtained, e.g. 1) land and 2) water, then four change comparator results are possible: 1) land to land, 2) land to water, 3) water to water, 4) water to land. An output product is made showing the subset of changes which is of interest to the user.

2. Delta Change Detection - This method is based on differencing (subtracting) remote sensor data from one time and another creating a new difference (called Delta) data set. Multispectral classification techniques are used to analyze and classify change in the Delta data. It is assumed that no change conditions will result in nominally zero values and will be combined into one general class. Only one classification is required with this technique rather than two for direct change detection.

3. Spectral/Temporal Change Classification - The aggregate of all spectral channels from the two dates is used in a change-no change classification in which the classifier is trained to separate out the different categories. Certain of the spectral/temporal classes would be due to all objects in the scene which did not change and others would be related to those which did change. This technique requires only one classification but it is an eight-channel classification.

4. Layered Spectral/Temporal Change Classification - This method is similar to (3) except that the layered classifier is used to perform a hierarchical classification from general classes down to specific classes. For example, change and no-change would be separated in one decision then further classifications would be performed to identify specific change.

Work has progressed on 1,2 and 4 and applications of 1 is furthest along. Direct change output has been produced for the Port O'Connor quadrangle using data collected on November 27, 1972 and December 15, 1973. The twenty class NS classifications produced by CY75 were generalized down to five broad, general classes and input to the detect change comparator. Of the twenty-five possible change classes eight to ten were selected as significant and outputs with these changes were produced and compared to maps and aerial photography. Many problems were encountered in evaluating the results, the three main ones being: location of individual pixels on maps and photos, lack of coincident photography, and tidal effects. These problems are being approached using overlays made from line printer printouts and interpretation overlays of actual scene cover all at 1:24,000 scale. A reference change map of each quadrangle studied will be generated to evaluate the results of the four approaches.

Some examples of initial test output are presented in Figure 1 for direct change detection in the Port O'Connor quadrangle. Image A and E in the figure are gray scale images using the .7-.8μm band to reproduce the general structure of the land and water areas in the quadrangle. Images B thru D and F thru H are individual change classes reproduced as white and all others as black. The result in D (water to sand) is a suspected tidal effect in which shore sand and channel spoil was exposed by lowered water level. Images F and H indicate a considerable variation in marsh location. These results are, however, based on the assumption of correct classification which may not be valid.

A difference (Delta) data set has been produced for the Port O'Connor quadrangle utilizing the November 27, 1972 and December 15, 1973 data. Analysis of the Delta data set is progressing. Selected areas have been clustered but interpretation of the results is proving difficult due to the lack of coincident ground truth.

The Layered Spectral/Temporal Changes Classification technique has also been applied to the November 27, 1972 - December 15, 1973 temporal data set. For this study, the ability of the layered classifier to use different feature sets at various stages of the decision tree is being exploited.

Statistics from date t_1 are used in the earlier stages of the tree, with the final decisions based on statistics from date t_2. Statistics for both dates are obtained from the same set of training fields.

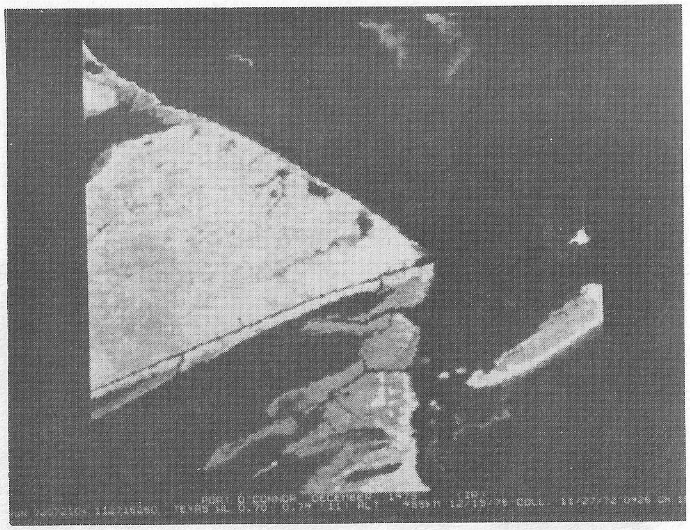
Figure 2 shows how a simple decision tree is designed to detect whether water changes to sand or remains water. Statistics from t_1 containing all classes are used to design the first tree. Statistics for the expected change classes are obtained from t_2 and a second tree is constructed. This tree is added on to the first tree at a node containing the class in which change is expected. Classification is carried out using the combined tree. The complexity of the tree increases as more classes are investigated for change and as the number of expected change classes increases. In a preliminary classification using a decision tree, similar to the one illustrated in Figure 2 but with more classes, changes of sandbars were detected off the coast of Port O'Connor.

Future work will complete the evaluation of all four techniques for at least two quadrangles over one change period. Later work will consist of a detailed evaluation of the "best" method for all quadrangles over a number of change periods.

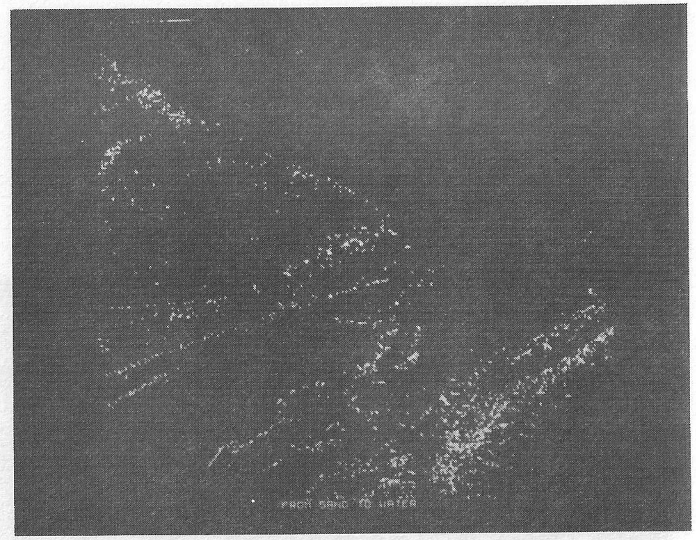
The major problem area in this effort is the lack of coincident photography or ground truth to correlate with the LANDSAT data. While the previously described data sets and their supporting data are probably sufficient to allow for technique development, they will not allow for a valid evaluation of the accuracy of any change detection procedure. It is suggested that an effort be undertaken to obtain another set of data similar to the February 25, 1975 data, (LANDSAT data with coincident color and color IR photography) and that this data be collected in late February 1976.

The results of this work would also be more useful to the ultimate users within the Texas State government if the changes of importance within the Coastal Zone could be defined. Presently, only gross changes such as marsh to land, water to land, land to marsh, land to water, etc. are being considered. These changes may or may not be of importance in the management of the Coastal Zone.

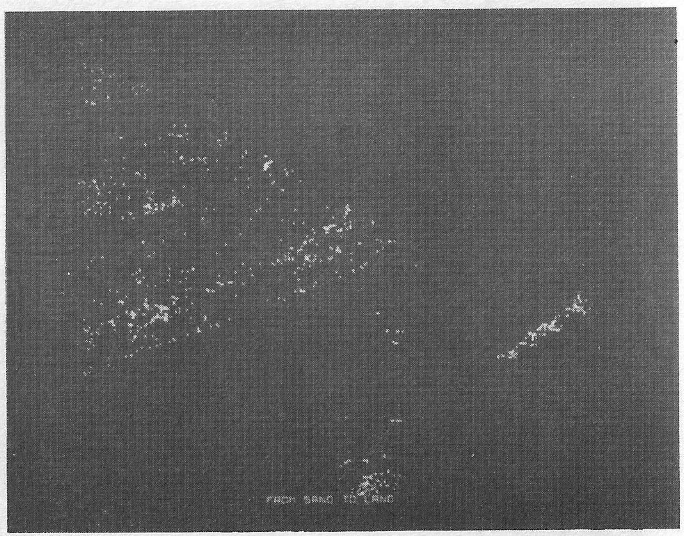
In addition to the change detection effort, each of the nine quadrangle areas involved in the CY75 effort was classified using a processor called ECHO (Extraction and Classification of Homogeneous Objects). This classifier is a per field classifier and initial results suggest that it may be more suitable for classifying the Coastal Zone than the per point classifier. Results of this work were forwarded to NASA/JSC for evaluation.



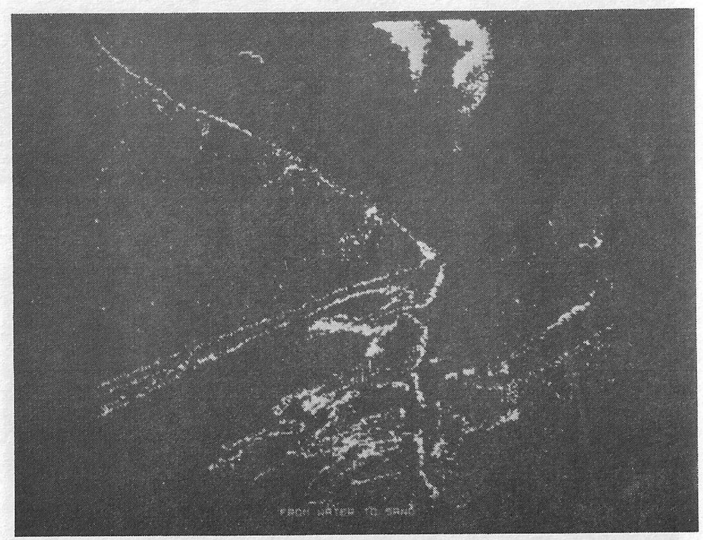
A. Reference



B. Sand to Water



C. Sand to Land

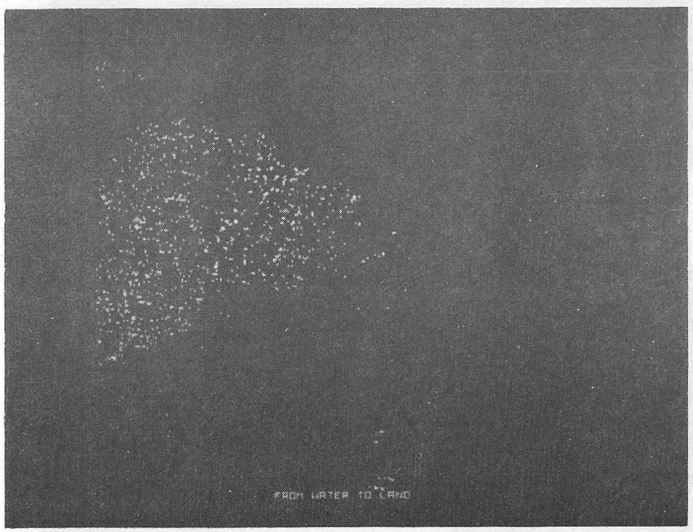


D. Water to Sand

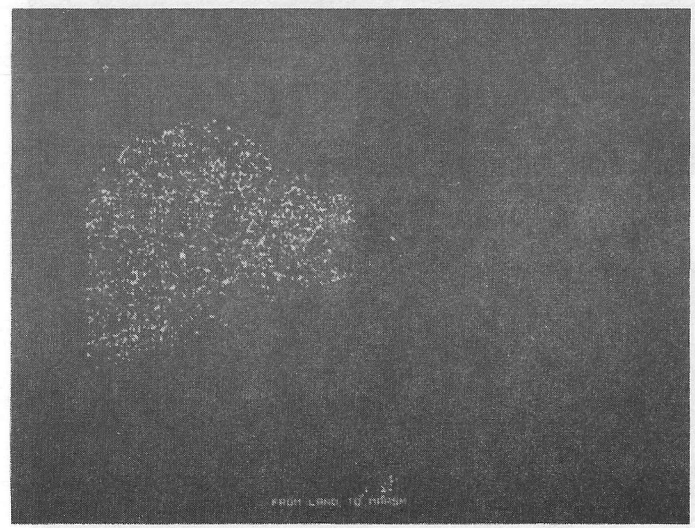
Figure 1. Direct Change Detection Results for the Port O'Conner Quadrangle for the November 1972 - December 1973 Period.

E. Reference

F. Marsh to Land



G. Water to Land



H. Land to Marsh

Figure 1. Continued

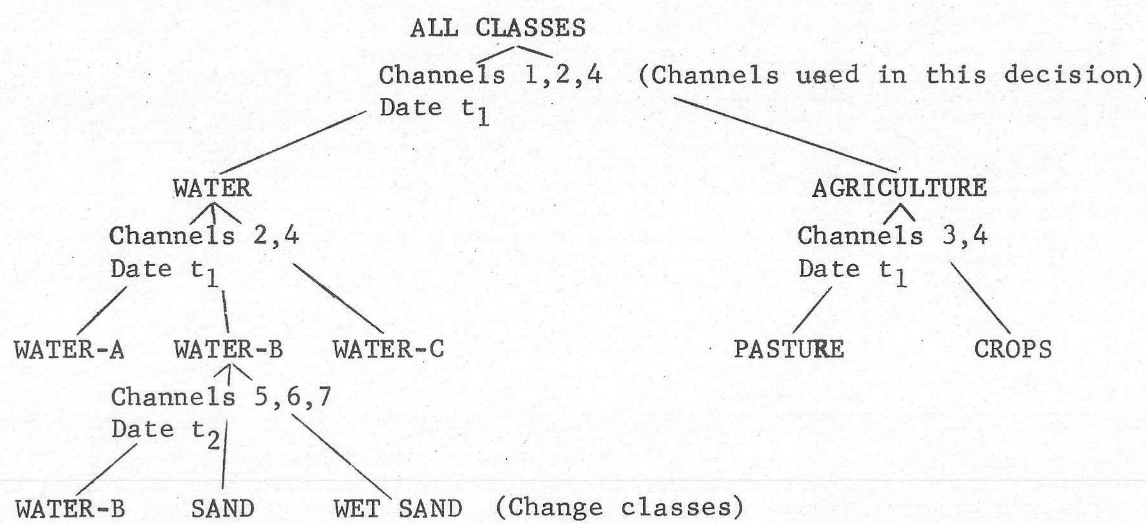


Figure 2. Decision Tree Design for Change Detection

Information Notes Issued
June 1, 1975 through December 31, 1975

103174 Anuta, Paul. Spline Function Approximation Techniques for Image Geometric Distortion Representations. Registration and geometric correction of multispectral scanner imagery from aircraft and satellite systems requires a mathematical representation to interpolate the image distortion between control points. For certain cases the distortions are severe and high order (eg., fourth, fifth or sixth order) polynomials are required which are difficult to handle computationally. In the work reported here, spline function approximation techniques which use a set of two-dimensional cubic polynomials joined with continuity in value and derivative are explored. Basic theory is presented and development of a two-dimensional cubic spline algorithm is discussed.

110574 Lindenlaub, John C. and Shirley M. Davis. LARSYS Educational Package: Instructor's Notes. The Instructor's Notes includes a "Survey of the LARSYS Educational Package" along with a flowchart showing the relationship of the seven units in the series. Relevant instructor information for all the units is contained along with copies of the student notes for Units II, III, IV and V.

050575 Cary, Tina K. and John C. Lindenlaub. A Case Study Using LARSYS for Analysis of LANDSAT Data. This material was designed to be a component of the LARSYS Educational Package, teaching new users the sequence of steps comprising an analysis. For each step of the sequence, concepts and theory are presented, an example using LANDSAT data is shown and discussed, and instructions are included directing the student to perform an analysis of another LANDSAT data set.

052875 Baker, J. R. and E. M. Mikhail. Geometric Analysis and Restitution of Digital Multispectral Scanner Data Arrays. This thesis contains the results of an investigation performed in order to define causes of geometric defects within digital multispectral scanner (MSS) data arrays, to analyze the nature of the resulting geometric errors, and to investigate restitution methods to correct or reduce such geometric errors.

The introduction of geometric transformation relationships for scanned data, from which collinearity equations for MSS may be derived, serves as the basis of parametric methods of analysis and restitution of MSS digital data arrays. The linearization of these collinearity equations is presented, including consideration of the functional assumptions made in order to model the stochastic changes in the exterior orientation of the sensor down the flight line.

A proposed system for the geometric analysis and restitution of MSS digital data arrays is introduced. This procedure is used to test the methods of analysis and restitution, utilizing actual MSS data arrays from two aircraft flights. The results of these tests indicate that the collinearity equations can yield acceptable results when untilized for the analysis and restitution of such arrays.

052975 Davis, Shirley M. The FOCUS Series 1975: A Collection of Single-Concept Remote Sensing Educational Materials. The FOCUS Series has been developed to present basic remote sensing concepts in a simple, concise way. Issues currently available are collected here so that more people may know of their existance.

061275 Swain, P. H., C. L. Wu, D. A. Landgrebe and H. Hauska. Layered Classification Techniques for Remote Sensing Applications. Layered classification offers several advantages over the very familiar single-stage approach. The single-stage method of pattern classification utilizes all available features in a single test which assigns the "unknown" to a category according to a specific decision strategy (such as the maximum likelihood strategy). The layered classifier classifies the "unknown" through a sequence of tests, each of which may be dependent on the outcome of previous tests. Although the layered classifier was originally investigated as a means of improving classification accuracy and efficiency, it has become apparent that in the context of remote sensing data analysis, other advantages also accrue due to many of the special characteristics of both the data and the applications pursued. This paper outlines the layered classifier method and discusses several of the diverse applications to which this approach is well suited.

062775 Phillips, T. L., H. L. Grams, J. C. Lindenlaub, S. K. Schwingendorf, P. H. Swain, and W. R. Simmons. Remote Terminal System Evaluation. The Laboratory for Applications of Remote Sensing at Purdue University has developed an earth resources data processing system which is being used by both LARS personnel and remote terminal users in part to determine the value of the system for training, technology transfer, and data processing. The results of Purdue's participation in this project are documented in this report. The facility has been used at seven separate sites and demonstrated to be a cost effective system for training personnel and technology transfer as well as meeting many data processing needs.

072175 Bauer, Marvin E., Tina K. Cary, Barbara J. Davis and Philip H. Swain. Crop Identification Technology Assessment for Remote Sensing (CITARS). This report summarizes the results of classifications and experiments performed by LARS/Purdue University for the Crop Identification Technology Assessment for Remote Sensing (CITARS) project.