LARS Information Note ~~110376~~ 091769

AUTOMATIC PROCESSING OF EARTH RESOURCES DATA

D. A. LANDGREBE

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

1976

33 p. (cover incl.)

Published in the Proceedings of the Second Annual Earth Resources Aircraft Program Status Review, NASA Manned Spacecraft Center, Houston, Texas. September 16-18, 1969. Pages 24-1 to 24-32.

2-P
MJX

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

SECOND ANNUAL

EARTH RESOURCES AIRCRAFT PROGRAM
STATUS REVIEW

VOLUME II

AGRICULTURE/FORESTRY, AND SENSOR STUDIES

Presented at the

NASA Manned Spacecraft Center
Houston, Texas

September 16 to 18, 1969

MANNED SPACECRAFT CENTER
HOUSTON, TEXAS

FOREWORD

On September 16, 17, and 18, a review of various aspects of the Earth Resources Program was held at the Manned Spacecraft Center, Houston, Texas. Particular emphasis was placed on the results of analysis of data obtained with the Manned Spacecraft Center and other aircraft which have contributed data to the program.

The review was arranged in conjunction with the Department of Interior, Department of Agriculture and the Department of the Navy. Attendees and participants at the meeting included program investigators, their immediate associates, and program representatives from the above named agencies and ESSA and NASA.

The review was divided into the disciplinary areas of Geology, Geography, Hydrology, Agriculture and Forestry, and Oceanography. An additional session was held on instrumentation. Program investigators presented the results of their work in each of these areas. The material presented is being published in three volumes:

Vol I — GEOLOGY AND GEOGRAPHY

Vol II — AGRICULTURE, FORESTRY, AND SENSOR STUDIES

Vol III — HYDROLOGY AND OCEANOGRAPHY

The review provided a current assessment of the program for both management and technical personnel. It is important to note that the material presented represented the current status on ongoing programs and consequently complete technical analyses will be available at a later date.

SECTION 24

AUTOMATIC PROCESSING OF EARTH RESOURCE DATA

D. A. Landgrebe

Introduction

This presentation together with the one following by R. B. MacDonald constitute a report of work at the Purdue University Laboratory for Agricultural Remote Sensing (LARS) during the last year. The objective of LARS is to help develop knowledge and techniques necessary to design an earth resource system utilizing aerospace platforms. This objective is pursued through studies in instrumentation and measurement, data processing, basic studies regarding the nature of earth surface materials, and considerations regarding the specific need for and use of information which can be generated by such a system.

It was decided early to concentrate efforts in the data processing area at Purdue on techniques for those tasks involving such large quantities of data that machine processing would be required. Following this, an initial approach using spectral data was selected for study because this type of data appeared to permit processors of adequate speed and small size. These decisions resulted in the multispectral scanner being the prime data source and this is assumed hereafter.

Due to the quantity of data which such a system will produce, one must not only develop algorithms suitable for analyzing the data and reducing it to useful information but one must also develop new techniques in manipulating, formating, and interfacing with these data.

The remainder of this discussion is devoted to relating what progress has been made in data processing studies at LARS during the past year. It is broken into six segments covering the following: (1) Improvements in LARSYS, the LARS programming system, (2) Initiation of a System Parameter Sensitivity Study, (3) The Digital Image Display System, (4) Data Registration, (5) A new data classification scheme, and (6) A boundary-finding algorithm for multispectral data.

The LARSYS Programming System

At a meeting held here one year ago a computer programming system called LARSYS was first discussed. Figure 1 shows the general block diagram of this system. It is seen that the system contains segments for data handling operations of aircraft data, aircraft data analysis, ground truth processing, and several activities. Figure 2 shows a flow diagram of the nanner in which this system is used in an aircraft data analysis task. It is seen that the system inputs are the aircraft multispectral scanner data and ground truth information. After appropriate processing by LARSYSAH and LARSYSGT respectively the data is available to the researcher. He is then in a position to carry out a dialogue with the portion of the system called LARSYSAA.

Figure 3 shows a more detailed diagram of LARSYSAA. The central portion of this system is a pattern recognition algorithm which permits the researcher to classify the data into one of several preestablished categories. There are several steps through which one must pass in utilizing this system. First, one must hypothesize the categories into which one wishes to divide the data. These categories may be plant species, soil types, geologic formations, or other type of earth surface cover of interest. The criteria for selecting the categories or classes are that they must be (a) of importance from an information value standpoint and (b) they must be statistically separable as a result of the data.

The second step is the selection of training samples or samples which typically describe each of the various classes. During the course of these two steps the statistics processor of LARSYSAA may be used to help. This processor can provide presentations of several statistics, histograms, and spectral plots of candidate classes of data.

Having tentatively established the classes and training for them one must decide upon what spectral bands are to be used in analyzing the data. The feature selection processor provides the researcher with an aid for this problem. Briefly, the feature selection processor implements an algorithm which determines the statistical distance between data sets in N-dimensional multispectral data space. Through calculations involving all possible N-tupples of spectral bands and all possible two-tupples of classes, one can decide not only upon the optimal set of spectral bands to be used but one can also get an initial determination of the degree of separatability of the classes.

The third processor, the classification processor, implements a maximum likelihood classification scheme and writes the result of the classification of each point on magnetic tape for further evaluation. This evaluation process is aided by the use of the fourth processor, the display processor which is capable of outputs giving both quantitative and qualitative information about the performance of the particular classification just carried out.

The general state of development of this whole technique can perhaps be summarized as follows. By the beginning of the 1969 crop year the system had been developed and demonstrated to the point of being able to adequately handle and classify data sets covering areas as large as a few square miles. The broad general objective for the 1969 crop year was to extend the size of area over which the system could be put to use to between tens and hundreds of square miles. The test site to be used for this development was Tippecanoe County in the State of Indiana, an area of 501 square miles. Flight lines were defined covering the length and breadth of the county and include a considerable percentage of the total land area of the county. Most of the flight lines are in excess of 20 miles in length. This is one to two orders of magnitude larger quantities of data then had been used previously and the question simply became can tecnhiques be determined for utilizing these methods successfully for quantities of data of that size.

As a result of this objective during the last year, a number of improvements have been incorporated into LARSYS. For example, an option was added to the feature selection processor such that in addition to using it to select the optimum feature set it can be used to group a large number of agricultural fields into those which are spectrally similar. This is useful in dealing with very large numbers of agricultural fields and in establishing the appropriate number of subcategories to use for a given crop type, for example.

In addition to modifications to the system for purpose of dealing with larger quantities of data, it was also recognized that the program needed to be made more general. At the time the programs were first written, the University of Michigan multispectral scanner system was the only system capable of supplying data for this type of algorithm. The programming had originally, therefore, been written specifically to process data from this scanner. The NASA scanner system placed under contract during the year has quite different and more general capabilities. Thus several modifications in the system were begun. Among these are the following:

*The system is being modified to handle any number of channels up to 30 and any number of samples per line.

*The calibration procedure used in the system is being considerably generalized so as to permit much more general tests and evaluation of calibration procedures.

*The system is being converted to work in terms of floating point numbers rather than fixed point numbers. This permits a more appropriate use of the dynamic range of the instruments among other things.

*Considerably more flexibility is provided to the user in that more functions can be performed and more parameters varied through the use of the control cards.

*The program has been rewritten to considerably ease modifications into the program.

One further comment may be in order about this programming system. The value of the system to the Earth Resources Program is considerably decreased if the system is only available to people carrying out research at Purdue University. As a result we have begun to give consideration to procedures by which the computer programs and the capability for data analysis which these programs make possible can be exported to other research centers. Since the programs exceed the equivalent of 15,000 Fortran cards, this is no small task. An initial step has been to temporarily fix the state of development of the system, document it at that stage, and transfer the computer programs to NASA/MSC. This has been completed in the last few weeks. I understand that NASA's purpose in this regard is to make sure that this type of capability is operationally available at MSC by the time the NASA scanner system is placed into operation.

In addition to the modifications indicated in the previous paragraphs this analysis system has been used extensively for data analysis purposes during the past year. Some of the results of these studies will be given by R. B. MacDonald in the next presentation. Further, LARS has recently been funded by NASA to determine the best way to apply these techniques to disciplines other than agriculture such as geology, hydrology, geography, etc.

System Parameter Study

Now that it seems clear pattern recognition techniques will be useful in the analysis of data in an operational earth resources system, there is a need to do a parametric study of the entire data stream to determine particular components of the system to which the analysis process is or may be particularly sensitive. Such a study would be beneficial in establishing adequately conservative but not overly restrictive (and therefore expensive) specifications for an operational system. One may reasonably ask questions such as the following:

What scanner sensor signal-to-noise ratio is really required?

What degrading effects occur in the processed data if the resolution of a single channel is considerably lower than other channels?

What are the effects of quantization noise on the analysis results?

In what manner do the analysis results degrade as the registration between spectral bands is degraded?

For the answers to most of these types of questions one can only speculate at this point as specific experimentation has not been carried out.

What we wish to do then is simply to indicate the need for such a parametric study and give some early results which have been obtained with regard to these questions.

First of all, consider the question of quantization precision. Data from the Michigan system is presently being quantized to 8 bit precision. This figure was selected based upon scanner signal-to-noise ratio considerations. During the year a data tape was constructed which contained 12 spectral bands as follows: ten spectral bands were recorded at 8 bit precision and two of the same 10 were also included but at a simulated 9 bit precision. In analyzing the data it was found that the feature selection algorithm favored these 9 bit bands over those of 8 bit precision from the same portion of the spectrum. The resulting classification accuracy was also slightly higher. Only a tentative conclusion should be drawn until these results are made more quantitative.

A question very often asked is how many spectral bands are required for a given classification? As one might expect, experimentation to date has shown that the answer to this question clearly depends upon what the categories are. In addition, consider Figure 4. This figure shows some results reported previously for another purpose with regard to the feature selection algorithm being used in LARSYSAA. The figure, as can be seen, is a plot of the percentage error as a function of the number of spectral bands used. The particular study for which these results were obtained was designed to show that the procedures used for band selection did indeed indicate the best possible sets of bands of a given size. That is to say, if as was the case in this study, one has a total of 12 spectral bands available and one wishes to use only three of the bands, which three will produce the best results.

The perhaps unexpected results which are apparent here however, are that not only is one able to pick the best subset of bands but that a subset of bands may produce a lower over all error than using the complete set of spectral bands.

At least using hindsight one may suspect that this phenomenon should occur; as seen from Figure 4 it has now been seen experimentally. It is also of interest to know if it can be predicted on a theoretical basis. I wish to show briefly here some results derived from a theoretical view-point and reported by Hughes in the IEEE Transaction of Information Theory, January 1968. These results are summarized in the next three figures. Figure 5 shows a graph of accuracy versus what Hughes calls measurement complexity. The assumptions here are that one is using a Bayesian classifier in a two class problem and that one has an infinite number of training samples with which to train the classifier. Further, the data is assumed to be in discrete (i.e., digital) form and the results are averaged over the ensemble of all possible pattern recognition problems.

By measurement complexity is meant in our case, the product of the number of possible discrete values that data in a given spectral band can produce times the number of spectral bands available. This figure shows what one would expect, that is that as the measurement complexity (which, keep in mind, includes the number of spectral bands) as this complexity increases the accuracy continues to increase indefinitely. But of course after some point, the curve becomes reasonably flat.

However, the infinite training sample set assumption does not fit our conditions. Hughes has also derived the results for finite numbers of training samples. This is given in Figure 6. Shown here is the case where each of the two classes is equally probable. One sees immediately that the curve increases as the measurement complexity is increased for awhile. It then reaches a maximum and begins to decrease with increased complexity. On the basis of this curve then one would reasonably expect if he had not seen it experimentally before that there would be an optimum number of spectral bands for a given signal-to-noise ratio.

There is one further case for which Hughes has obtained results which are of interest to us. This is the case when the classes are not equally probable. Consider Figure 7. These are the results when one class has a probability of occurence of .2. Since Hughes has assumed a two class situation the other class, therefore, has a probability of occurence of .8. Notice from this curve that there is not only a maximum but also a minimum indicating that other things being equal there is not only a best number of bands to use but there may also be a worst number of bands.

We do not wish to imply here that these results apply directly to any specific remote sensing problem and to the classification of multi-spectral data in the multi-class environment since the assumptions in Hughes work are somewhat different. Nevertheless, the two problems seem similar enough that a study of the effects of the data signal-to-noise ratio (we assume here that the data is digitized to a precision based upon its signal-to-noise ratio) as well as the number of spectral bands is indicated.

Digital Image Display System

Most of the discussions up to this point have been with regard to work in software systems. There has recently been completed however, the design for a major piece of hardware. I refer to the digital image display system on which work was begun some three years ago. At that time it was recognized that when interfacing with very large quantities of data one of the chief problems would be the matter of editing out those specific parts of the data with which one wished to be concerned. There appeared to be no suitable piece of hardware by which to achieve this capability and it was at that point that use of the now-familiar computer line printer printout technique as shown in Figure 8 was begun. This figure shows a conventional aerial photograph and a computer-line printer generated printout of a single channel of a multispectral data taken at the same time. The particular printout shown here happens to have 440 symbols per line, and 12 different symbols simulating a 12 step gray scale has been used. This kind of scheme can be used for data editing simply by locating the number of the scan line and the number of the sample within the scan line after viewing the printout.

This scheme works very well in many applications, particularly those where more limited quantities of data are involved. However, it does bog down when the quantity of data becomes very large simply because of the sheer amount of paper involved in the printout. Thus three years ago the design of the digital display system was begun as an answer to this problem.

Since that time it has become apparent that there are many more tasks besides simply data editing which require a human to be able to rapidly interface with a very large quantity of data. Thus the original design has been modified continually until it became essentially complete approximately a year ago. The acquisition for this equipment has been funded by NASA through the USDA and recently placed under contract for construction. Figure 9 shows a sketch of the display console itself. It consists of a display screen on which the imagery can be displayed utilizing 768 sample points per line and a gray scale of 16 steps. There is also provided a light pen with which the operator can designate to a computer a particular data point of interest and a keyboard consisting of 32 keys and as many as 256 keyboard overlays by which the computer operator can control the computer processing and the display of the data. The keyboard overlays (see Figure 10) contain sensor elements shown at the upper edge such that when the overlay is placed in position on the keyboard the computer can automatically sense which overlay is present and therefore which computer programs are to be called by each of the keys.

Very briefly the complete display system consists of a CRT system similar to a TV monitor and a disc buffer system. On signal the data is transferred from the computer interface through a core buffer to the disc buffer system. The disc system permits the screen to be continually refreshed at a rate adequate to prevent flicker on the screen. A photo-copy unit is also provided to produce output in hard copy form.

Data Overlay

For several years now we have been pursing studies directed toward the precision overlay or registration of one scanner image upon another. The basic problem is illustrated in Figure 11. Here we have scanner images from two different scanners, in this case a visual and a thermal image. An airphoto is included for clairty. One wishes to bring the two images into registration or in other words overlay one image upon the other in such a way that corresponding points in the two are aligned in every case.

Last year at this meeting we reported a procedure which was used reasonably successfully for this purpose. This procedure is illustrated in Figure 12. The basic process necessary is to numerically correlate scan lines from the two images under the assumption that registration is correct when the correlation is maximized.

However, it has been found that particularly in the case of widely different parts of the spectrum correlation of the scanlines directly did not provide good results. The correlation between images from very different parts of the spectrum is not sufficiently peaked. This difficulty was overcome by first applying a simple border enhancement algorithm to both images, then correlating the result. The final step is the writing of a new data tape upon which the data from the separate tapes (i.e., images) has been merged and properly aligned.

Figure 13 shows the result of the border enhancement on the two images for a certain data set. The border enhancement scheme used was simply to take the magnitude of the first difference adjacent points.

During this last year several improvements have been incorporated into this system. First of all two deminsional correlation is now being used. In addition as would be expected, in using the system it became immediately apparent that the effectiveness of the system is data, i.e., scene, dependent. Two pictures with a grid square pattern will be easier to correlate then, for example, two pictures of a smooth sea. One of the parameters available with regard to correlation difficulty is the data set size or window size used in the correlation. Too large a window unnecessarily slows the processing while one which is too small results in poor performance. In short what was needed was to make the system adaptive to the scene being processed.

This has been accomplished by defining and utilizing a "Picture Complexity Index" or PCI. The one used is defined in terms of the average number of border or boundary points in a cell of fixed size. To illustrate this consider Figure 14. This figure shows samples of three classes of imagery. The first class referred to as "Rectangular," would be relatively easy to correlate due to the number of distinct linear features. The second, referred to as "Natural" is relatively difficult to correlate and would require a larger correlation window due to the essential randomness of the structure. The third is intermediate to these two and is referred to as "Mixed."

Figure 15 shows border density histograms for these three scenes. This is merely a comparative plot of the number of 10 x 10 cells versus the number of border points per cell for each image. One would expect to find that a simple scene, in this case the one called "Rectangular," would have relatively more cells with a low number of border points and therefore the PCI which is the average number of border points per cell would be low. The opposite should be true for the "Natural" scene with the "Mixed" case intermediate. This indeed turned out to be the case. The resulting PCI's are given in the figure.

This "Picture Complexity Index" has now been incorporated into the overlay system as shown in Figure 16. As seen here after routine data handling operations, the border enhancement procedure is carried out and the PCI determined. This is used to set the range of a variable range correlation. This picture alignment operations complete the procedure. Ultimately, one of the most important uses of the overlay procedure will be in the overlay of the data gathered at different times of the day or year. This is a goal of the development.

Per Field Classification

The classification algorithm used in LARSYSAA utilized only spectral information and classifies points entirely without regard to the classification of neighboring points. As previously mentioned the reason behind this approach is the desirability of possible processing speed and simplicity. However, accuracy can certainly be increased by incorporating spatial information as well. Given that a certain point is a member of a certain class there is considerable likelihood that its neighbor is also.

With this in mind a new classification scheme, referred to as "Per Field Classification," has been proposed and is now being studied. Figure 17 shows an airphoto, gray scale printout of one band and a per point classification printout for the task wheat versus everything else. Suppose prior to classification one were to point to the second wheat field from the top of the gray scale printout and say, "I don't know to which of the classes this field belongs but whichever it is, all points in the field belong to the same class."

In this case one sees the possibility of a classification scheme which is fundamentally different than the per point classifier used in LARSYSAA. In the per point classifier one is comparing a single point with each of several conditional density functions (i.e., the class training samples) for membership likelihood. In the per field scheme one compares a point set or density function with the conditional densities. Intuitively one would expect the latter to be more accurate.

To see if this turns out to be the case three flight lines were selected from which data had been gathered during a time of the growing season for which analysis had proven difficult. Classes and training samples were chosen for each flight line. After choosing a set of spectral bands the flight lines were classified using LARSYSAA. Classification accuracies were determined in each case using sets of test fields which were as large as possible without using any of the training samples. Further, since these results were to be compared with the per field results, the accuracy was also tabulated by fields by assuming that if any field had 60 percent or more of its points assigned to a single class the entire field was considered correctly classified.

Following this the per field classification scheme was tested on the same data using the same training statistics, spectral bandsand test fields. The results are shown in the next three figures. Figure 18 shows the comparative results for a four class three band test of July 1966 flight line C-3. It is seen that the overall accuracy is increased from 77 percent to 91 percent using the per field scheme. A similar overall improvement resulted in July 1966 flight line C-4 data as seen in Figure 19. Figure 20 shows the results for September 1966 flight line C-2. We note by the way and in passing that though the term field might refer to an agricultural field, it need not; any set of points however defined which are presumed a priori to be from the same class can be classified by this scheme.

A Boundary-Finding Algorithm

As described above the per field classifier must be classed as a semi-automatic rather than a fully-automatic technique since it would apparently be necessary to manually delineate the field boundaries. One can imagine many circumstances under which it would be most useful as such. However, it can also be made fully-automatic if one can find a suitable method of automatically drawing boundaries between spectrally different surface cover types.

There are in remote sensing many other points at which a boundary-finding algorithm would be useful. We found the need to do this earlier in the overlay problem, for example. We have as a result begun to work on developing such a technique. At the outset this problem seemed to call for an unsupervised classification technique. A clustering algorithm was defind for this purpose and program logic suitable to image analysis of this type was established utilizing it. Basically the logic calls for the clustering of data in each cell of 9 x 9 points. Adjacent points which become assigned to different clusterpoints become delineated as boundary points.

Figure 21 shows an airphoto of a flight line, a gray scale printout and the result of using the boundary-finding algorithm on the same data. It is seen that a large number of the actual boundaries in the data are indeed located. There is in the algorithm logic an option to indicate the intensity of the boundary through control of a threshold. Figure 22 shows the result of changing the threshold so as to change the number of boundaries located.

This algorithm is very much in the early stages of its development and much work remains. For example for some applications it will be necessary to increase the sophistication of the program logic at least to the extent of being able to draw closed contours.

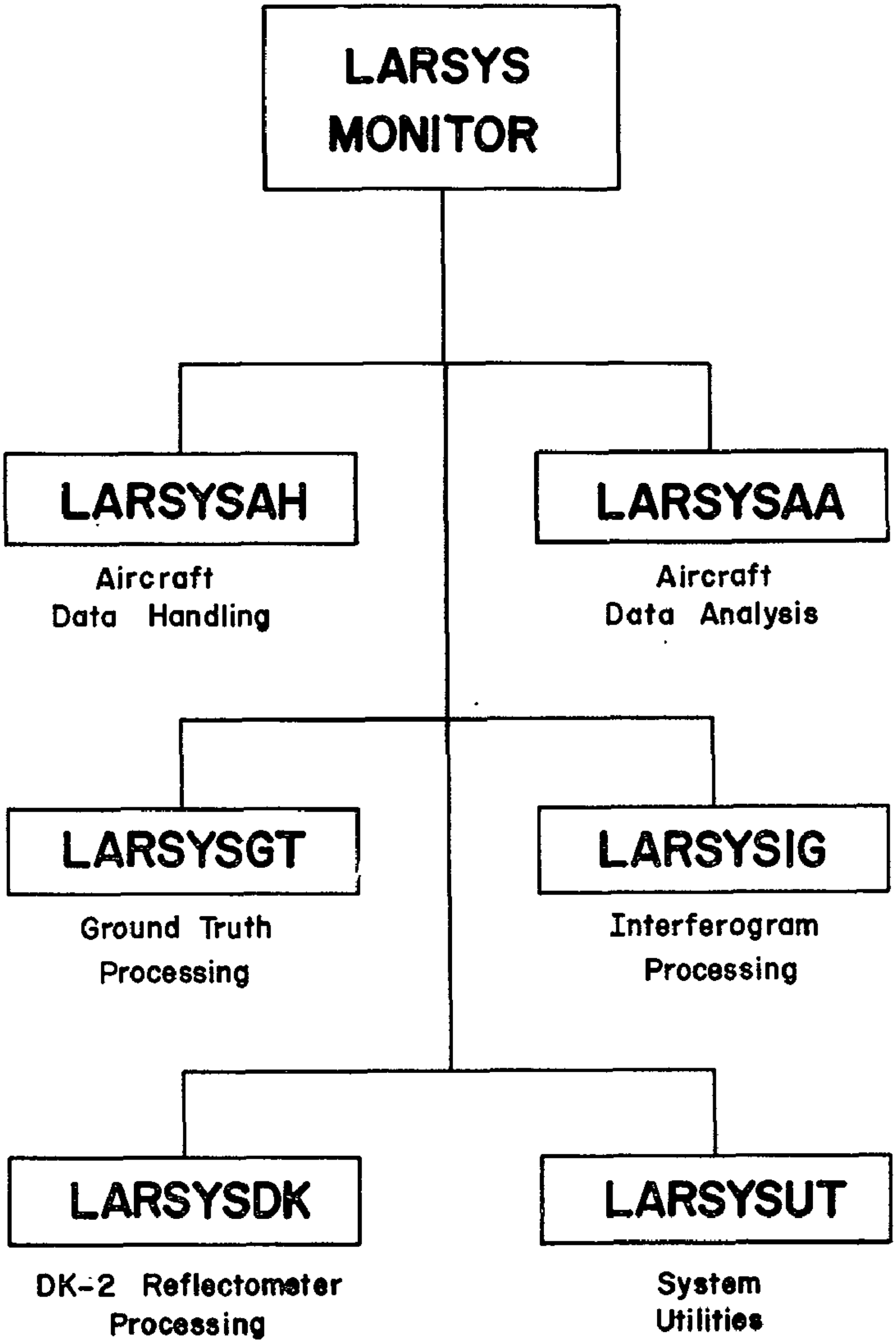


Figure 24-1.- Organization diagram of LARSYS programing system.

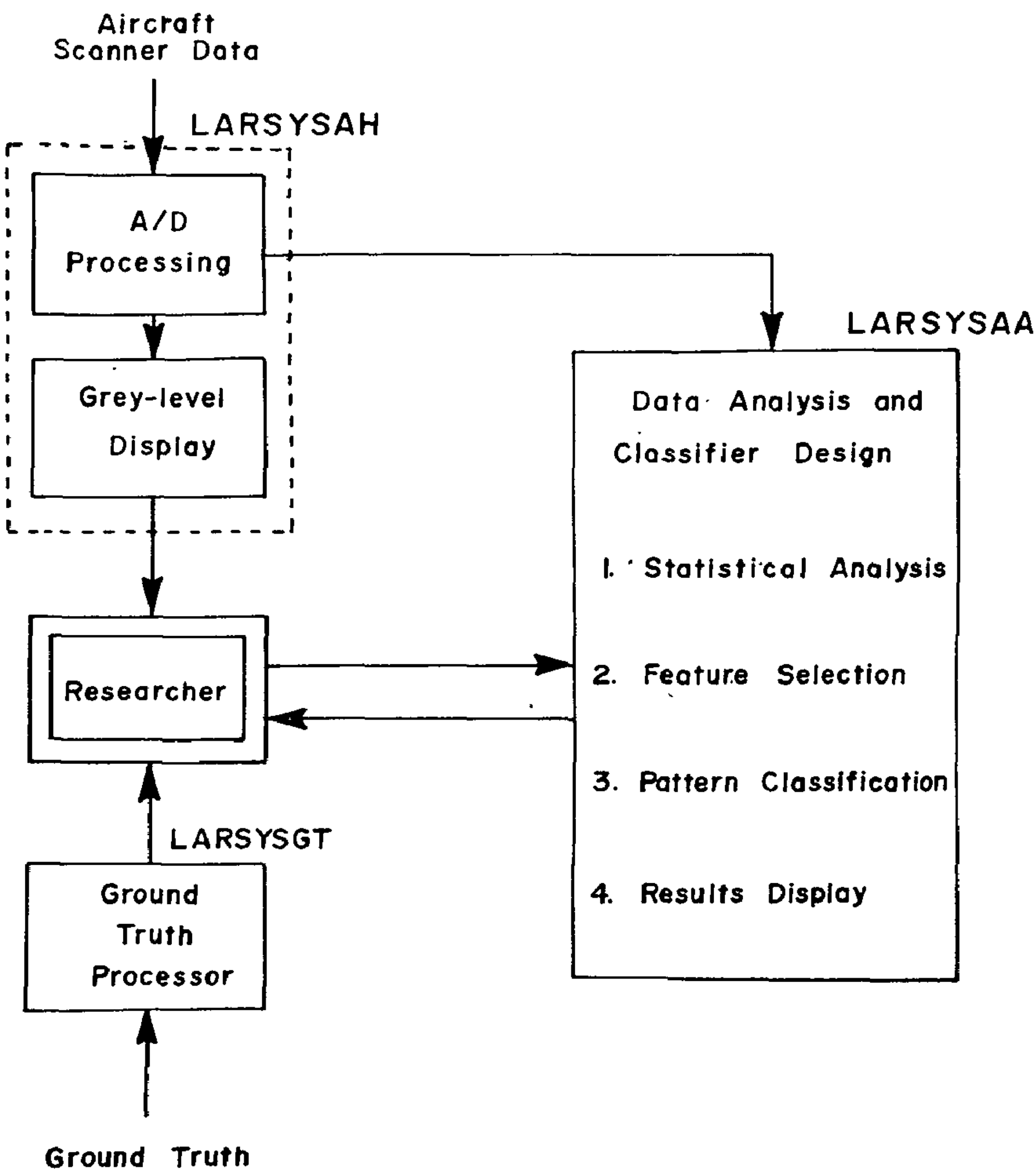


Figure 24-2.- Flow diagram of LARSYS system use for data classification studies.

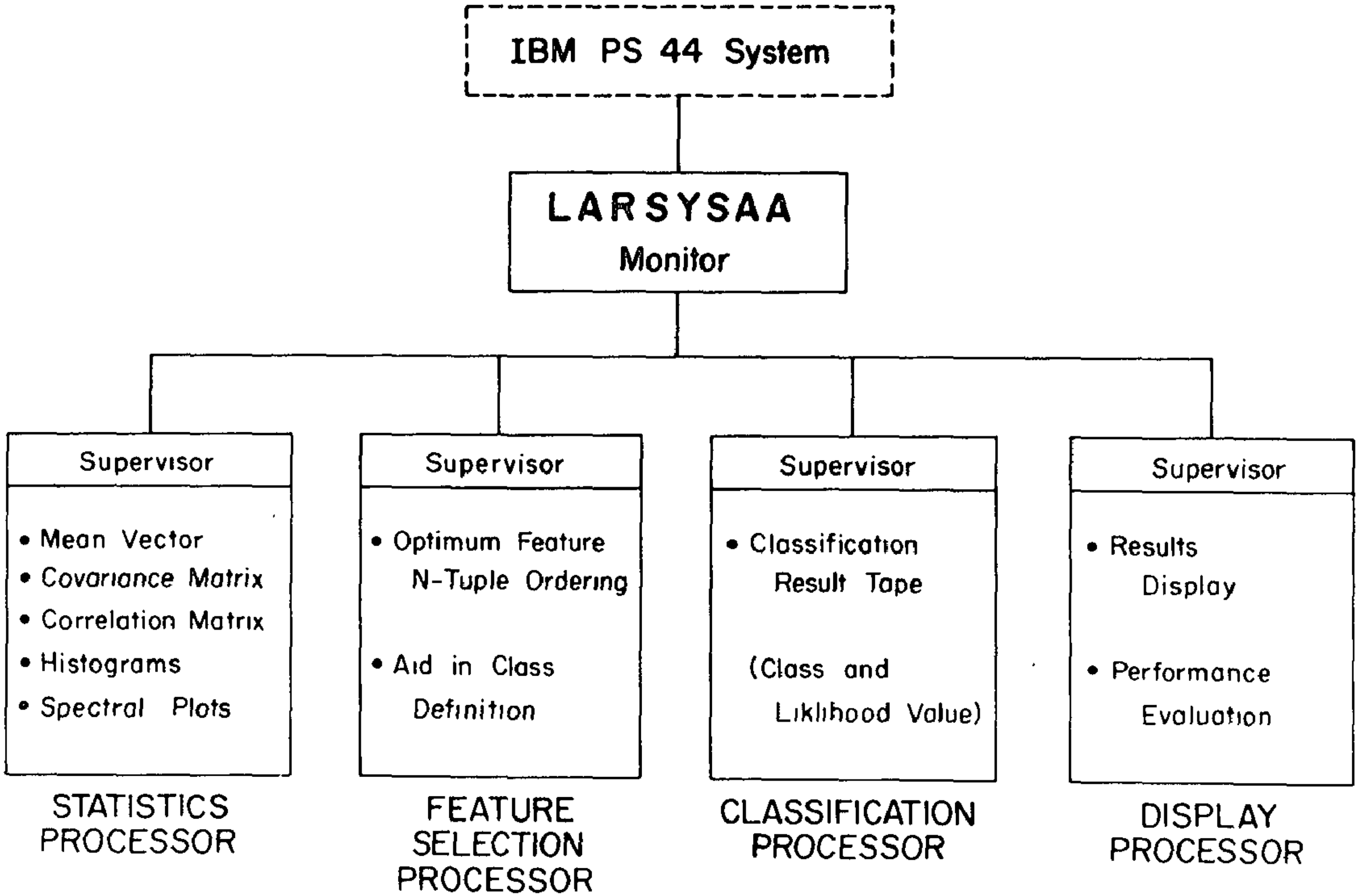


Figure 24-3.- Organizational diagram of LARSYSAA. This system enables the researcher to design and evaluate a pattern classifier.

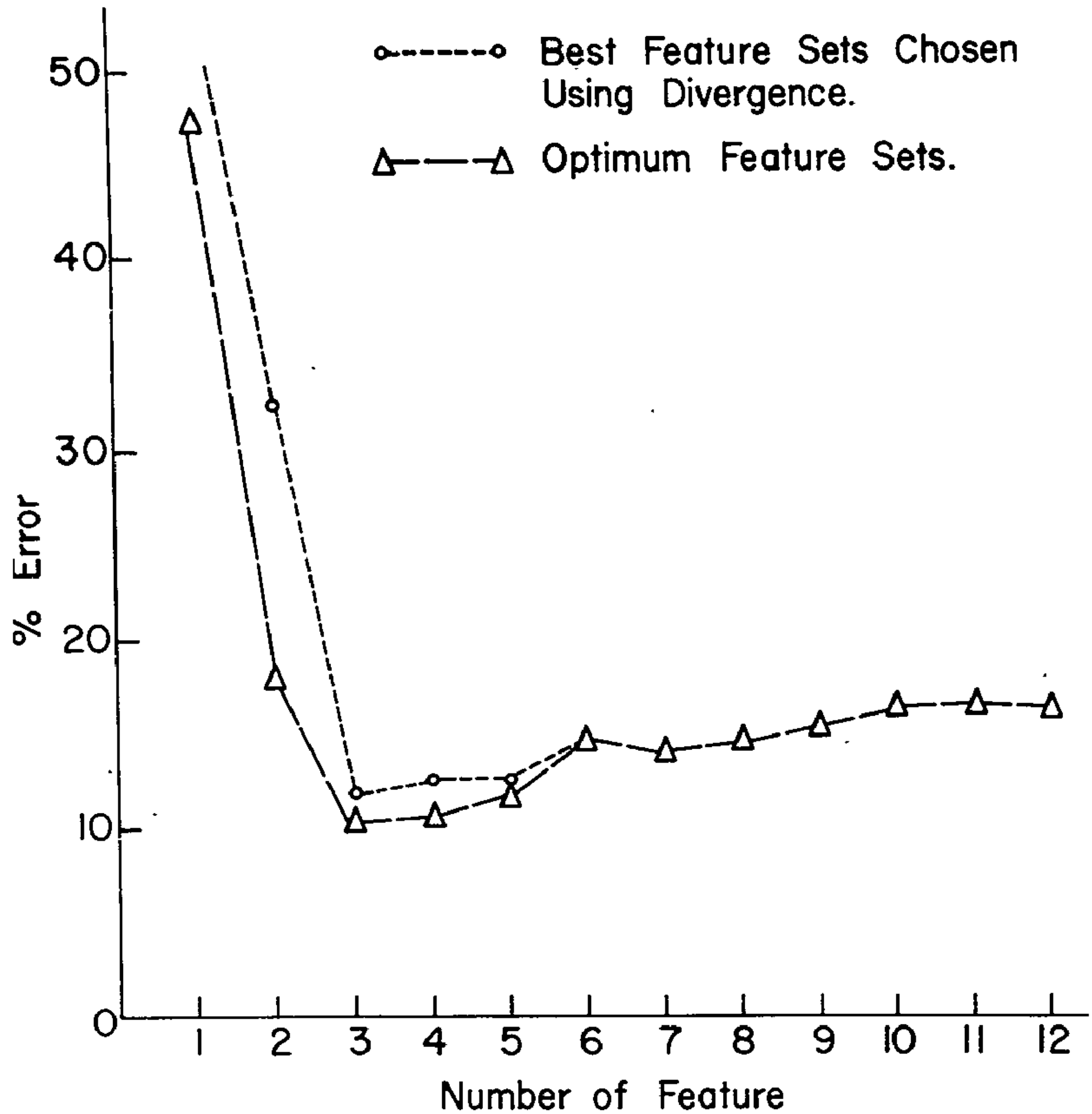


Figure 24-4.- Per cent error vs number of features (spectral bands) for a specific classification task. Note that better classification accuracy is obtained in this case using only three spectral bands than any other number.

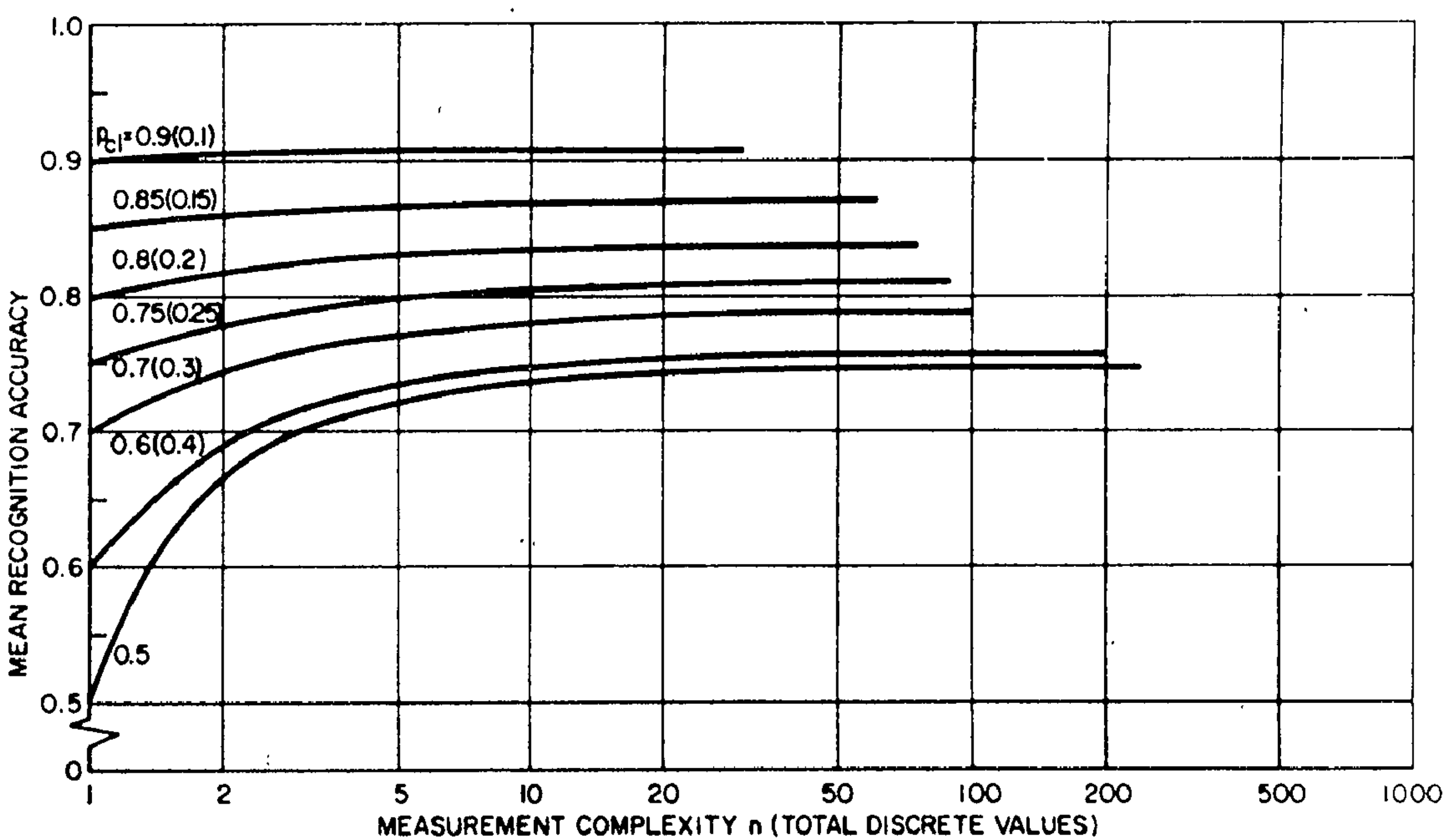


Figure 24-5.- Mean recognition accuracy vs measurement complexity for a training set infinite in size. The parameter of the graph is a priori class probability (Hughes, 1968).

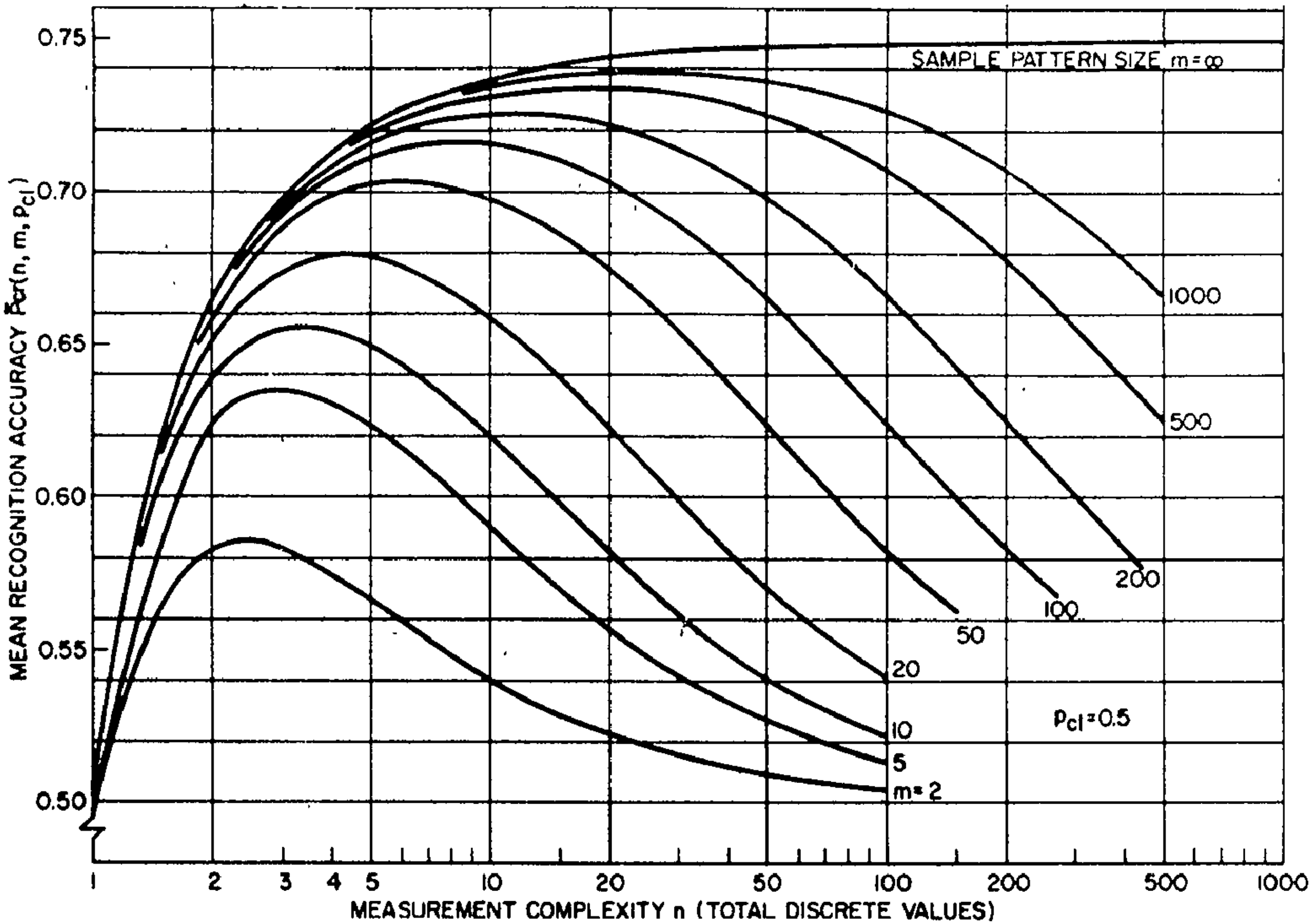


Figure 24-6.- Mean recognition accuracy vs measurement complexity for equal a priori class probability. The parameter m is the training set size (Hughes, 1968).

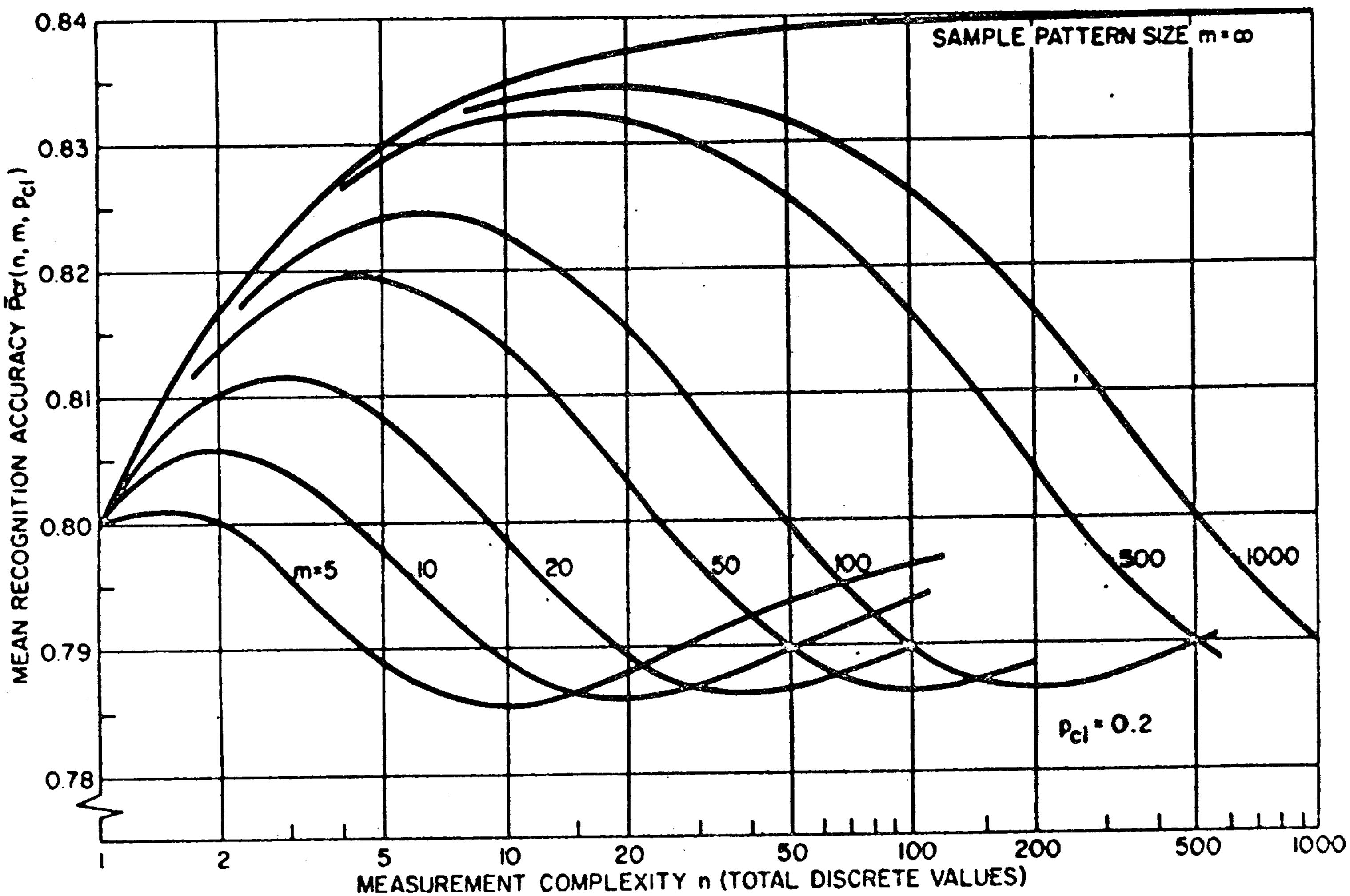
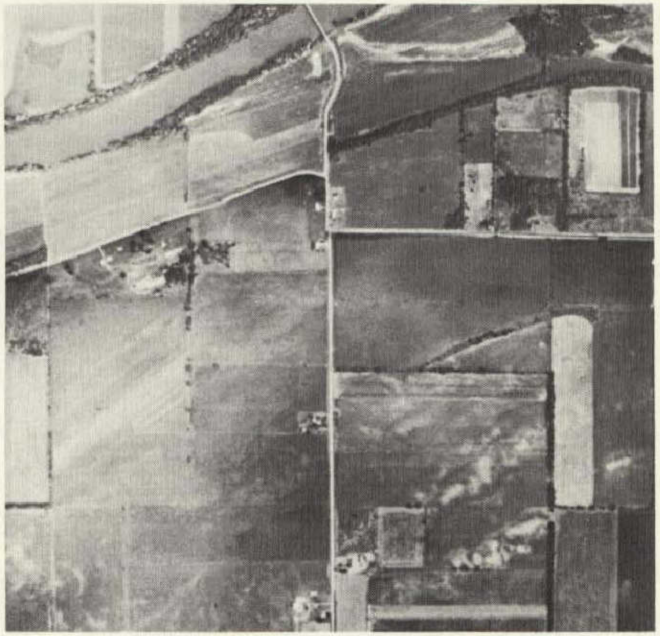


Figure 24-7.- Mean recognition accuracy vs measurement complexity for a priori class probability. The parameter m is the training set size (Hughes, 1968).



Panchromatic Aerial Photograph
(.4 - .7 micron)

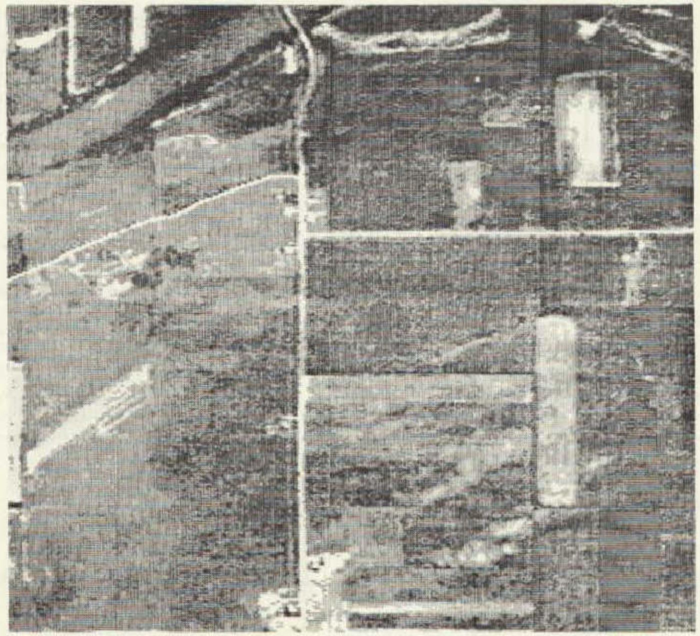


Image Simulated on Line Printer

Figure 24-8.- Comparison of a panchromatic aerial photograph and computer line printer simulation of scanner image from the .62-.66 micromter band.

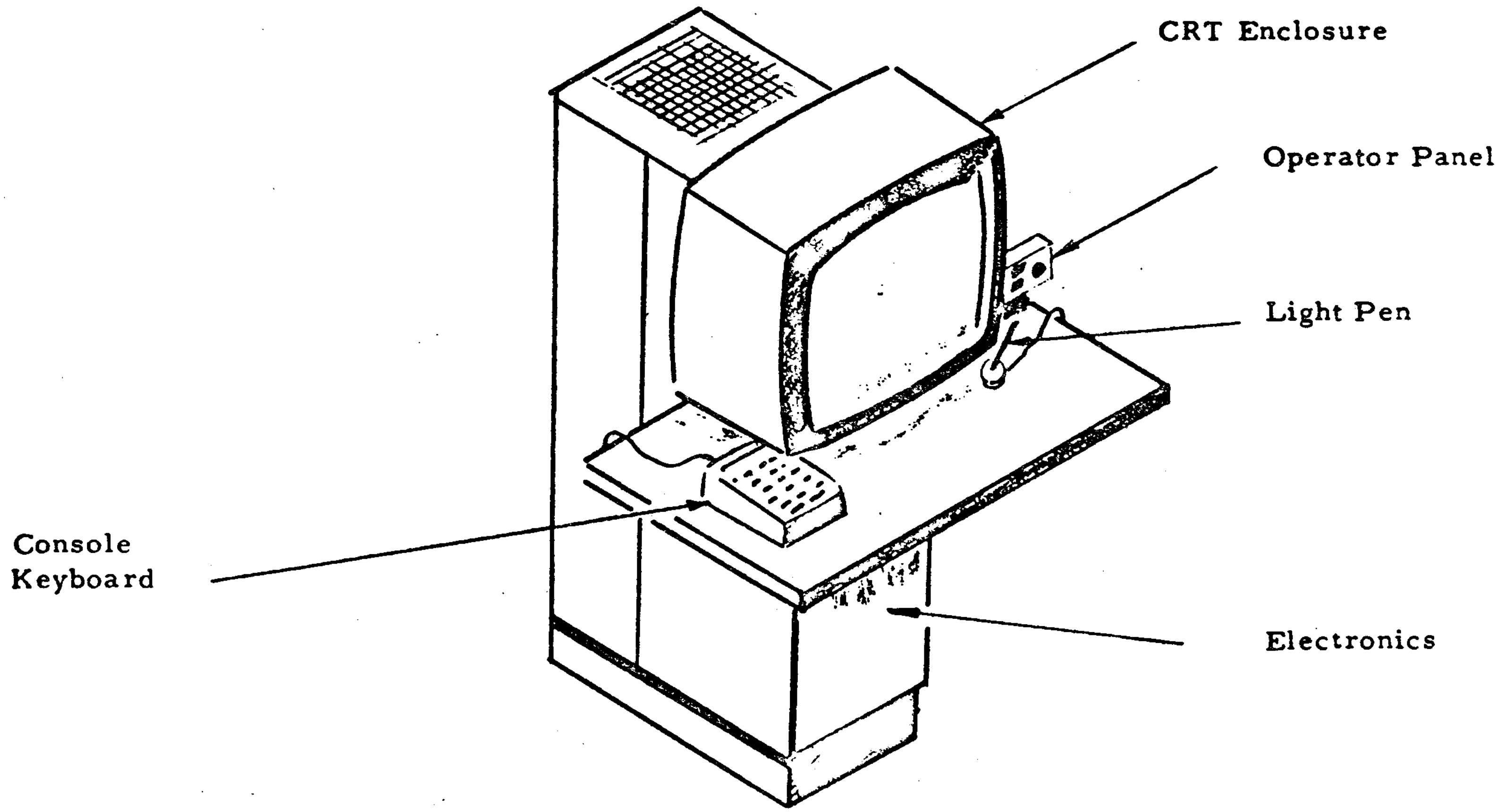


Figure 24-9.- Sketch of the Digital Image Display Edit Console.

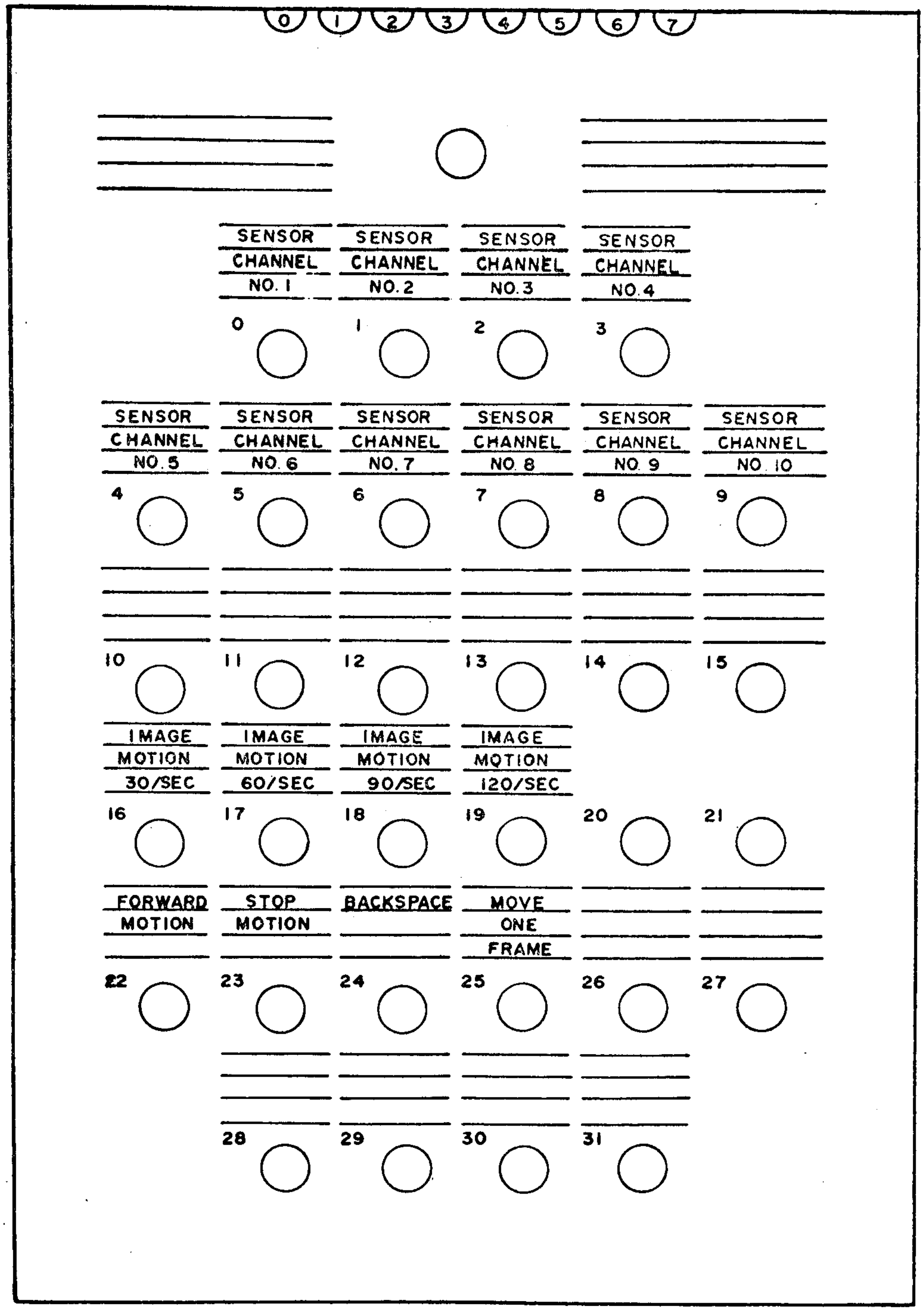


Figure 24-10.- Function keyboard overlay for the Digital Image Display System. Some sample keyboard functions are indicated on some of the keys.

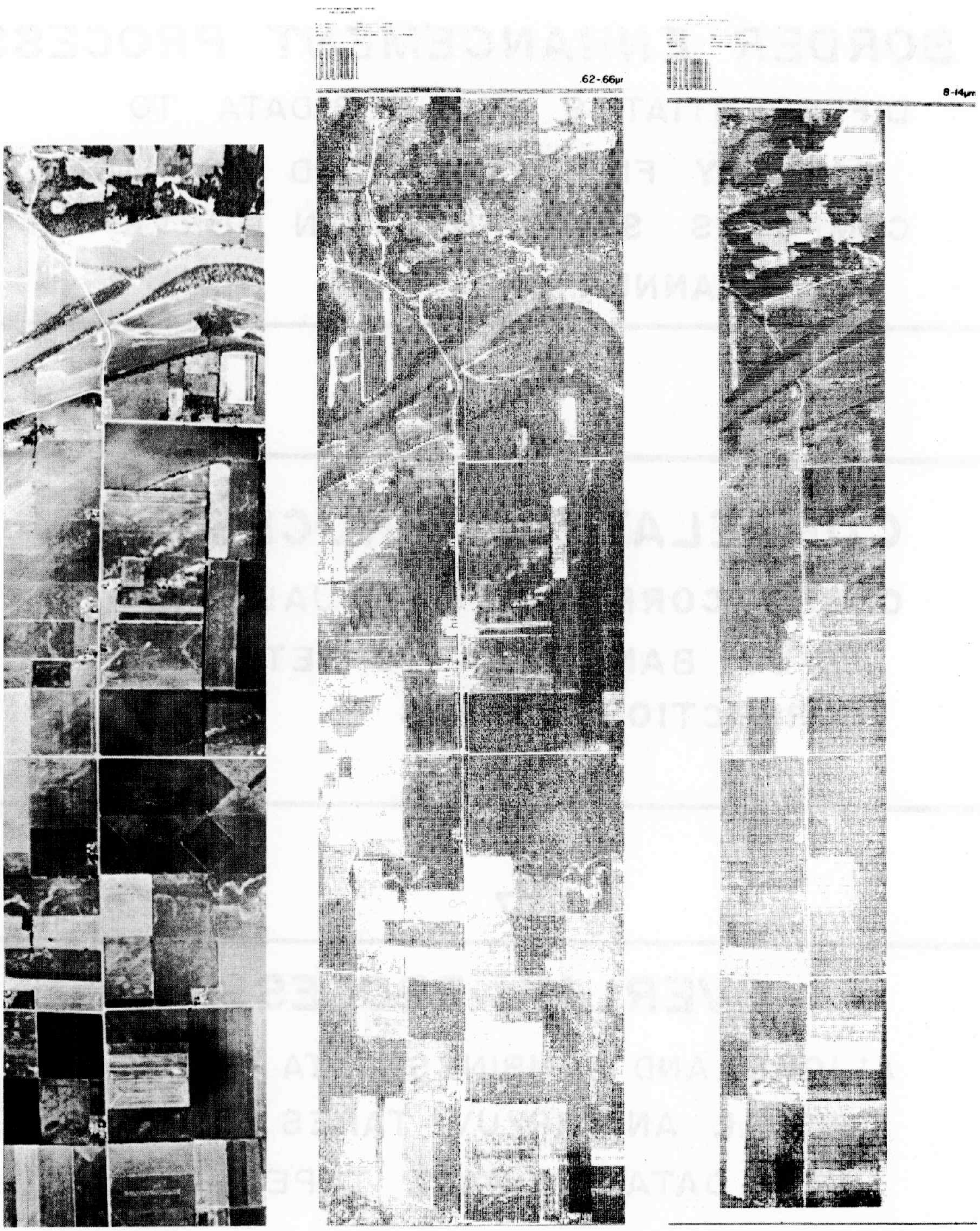


Figure 24-11.- Scanner imagery from a visible (.62-.66μm) and a thermal (8-14μm) band. A panchromatic airphoto is shown on the left for comparison.

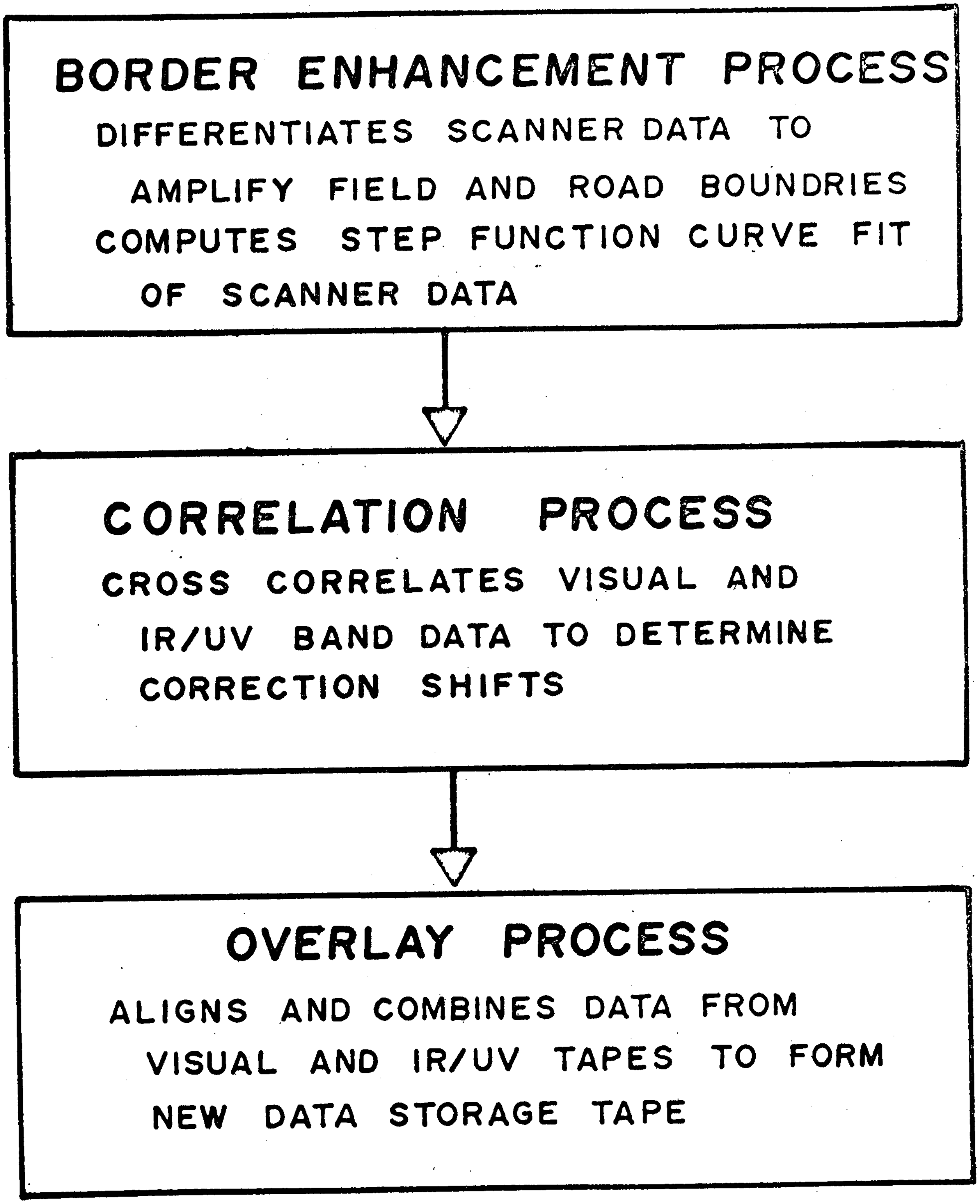


Figure 24-12.- Original organization of data overlay system.

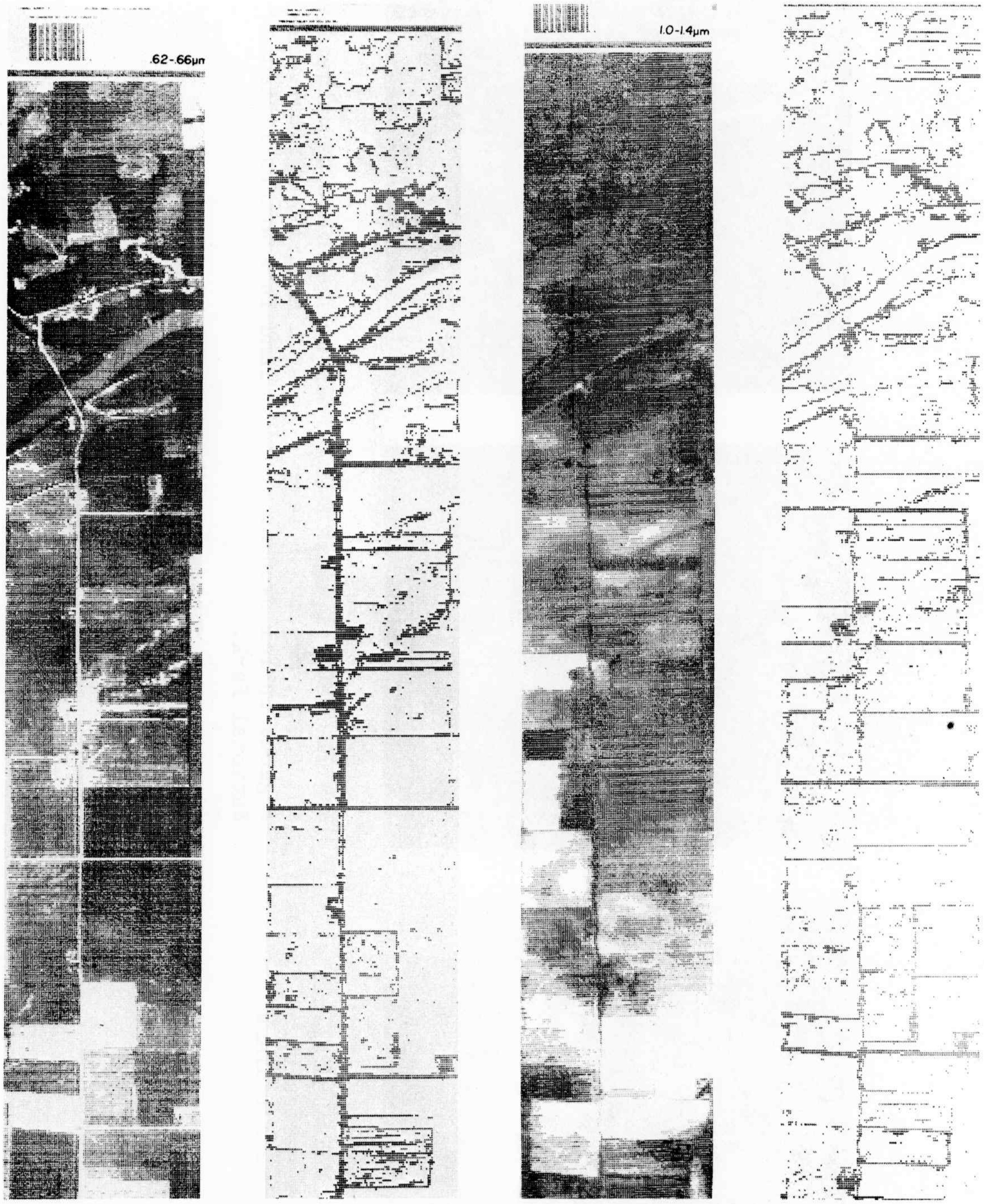
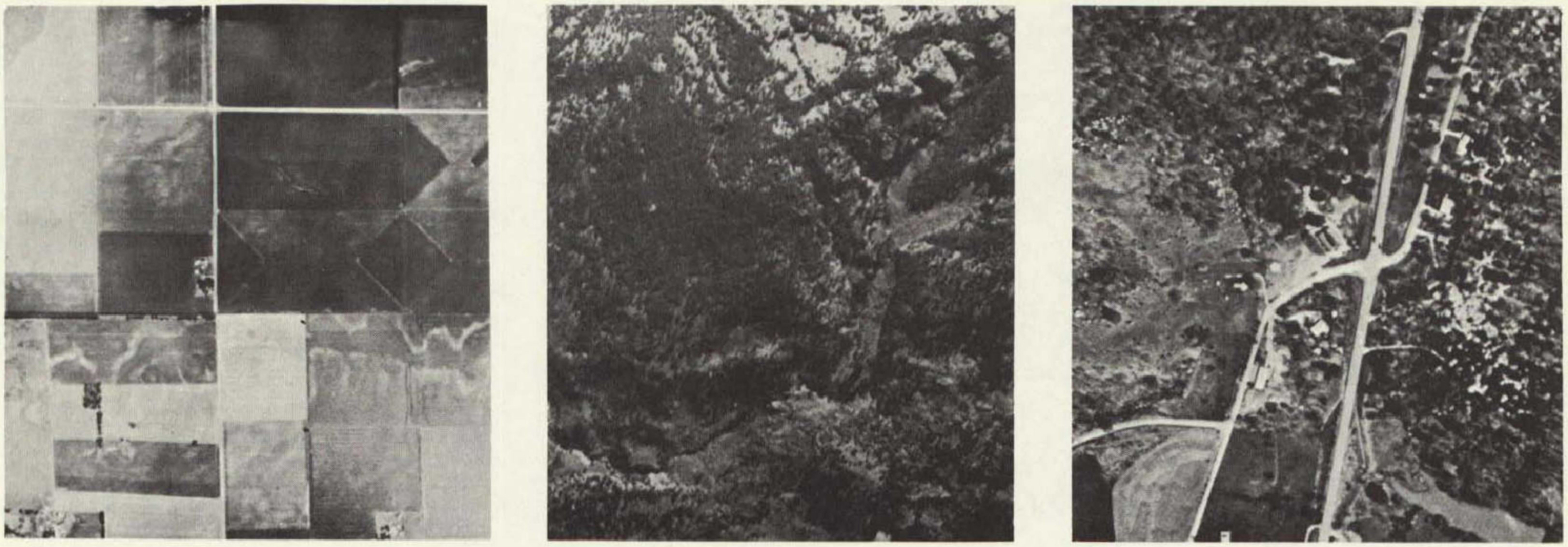


Figure 24-13.- Results of border enhancement on visable and thermal scanner images. A gray scale printout of each is also shown for comparison.



Class I- Rectangular plots such as agricultural fields in northern Indiana.

Class II- Natural areas such as forest, meadow or lake regions in Yellowstone National Park.

Class III- Mixed terrain cover such as hilly agricultural land in southern Indiana.

Figure 24-14.- Examples of three types of terrain cover.

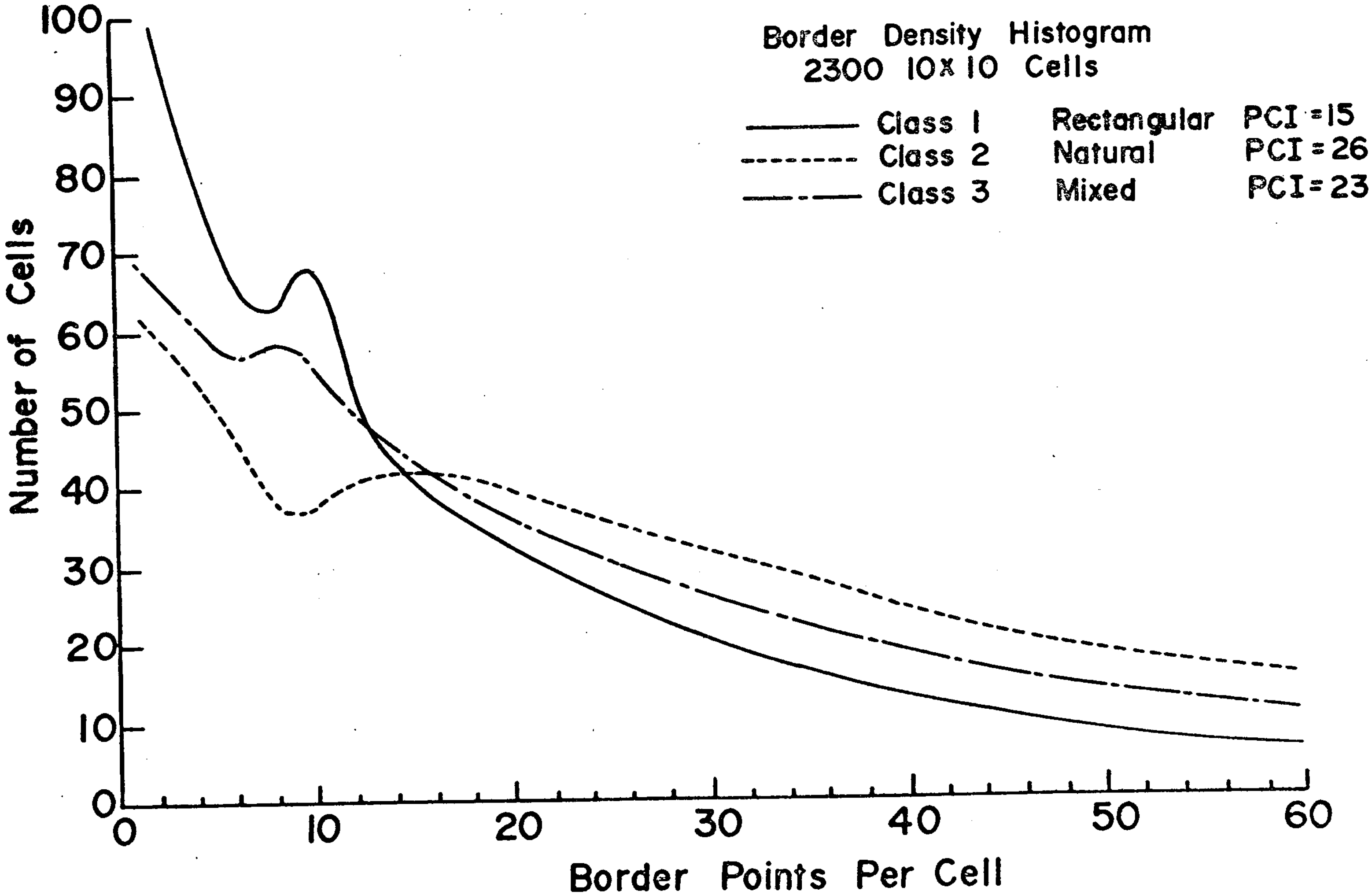


Figure 24-15.- Border density histograms for three classes of scenes.

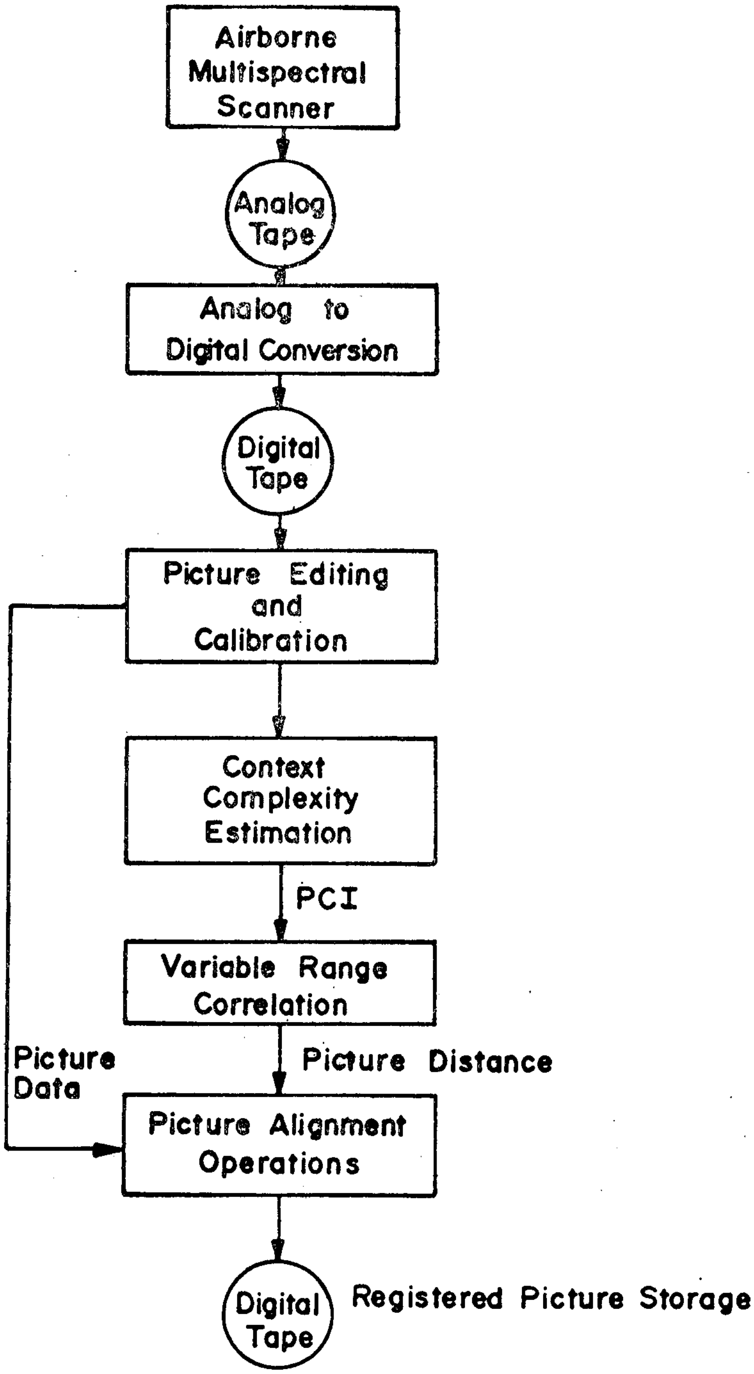
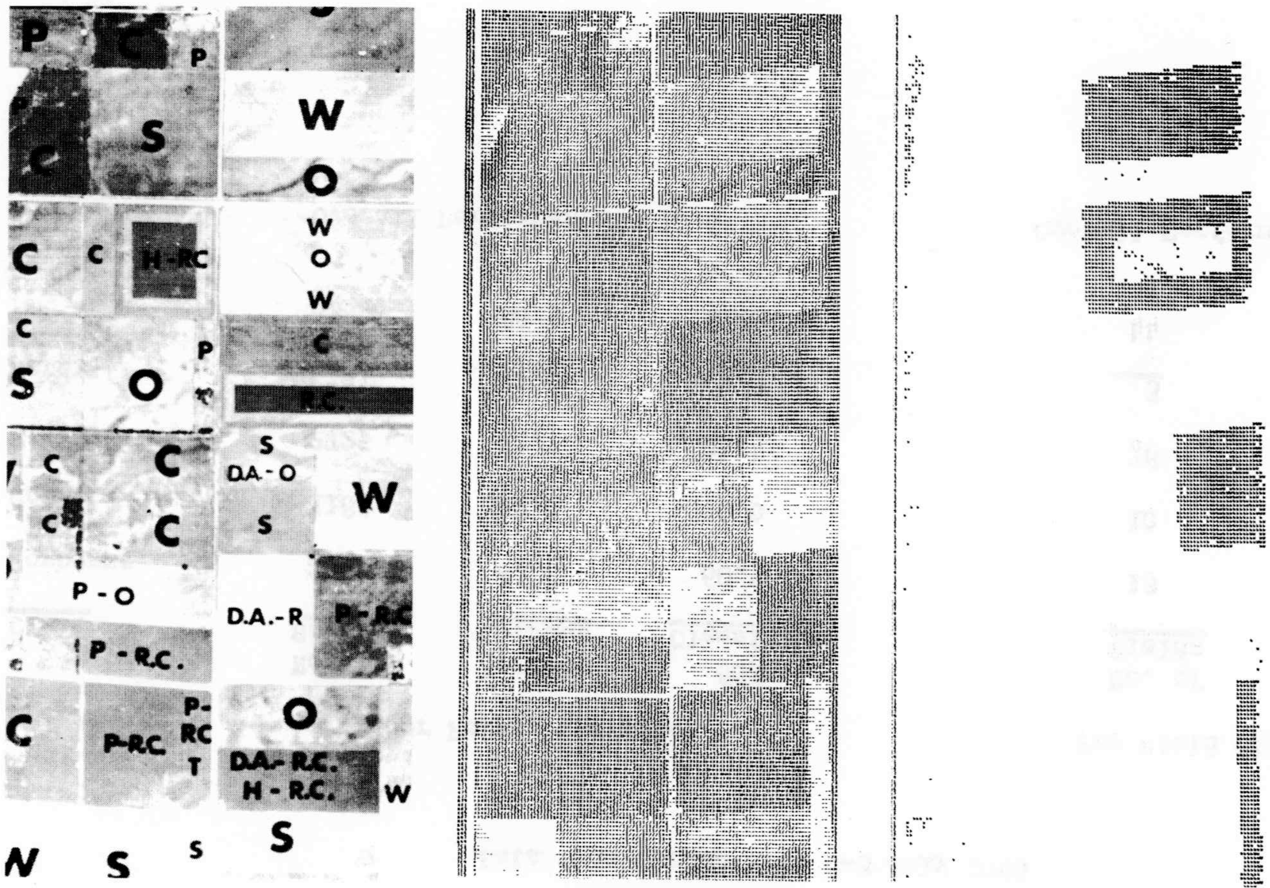


Figure 24-16.- Organization diagram of current data overlay system showing the adaptive feature using the picture complexity index.



a. b. c.

Figure 24-17.- Airphoto (a) gray scale printout (b) and LARSYSAA classification result (c) of a section of data for the classes wheat and other. The ground truth is indicated on the airphoto by symbols as follows: W - wheat, S - soybeans, C - corn, O - oats, R.C. - red clover, R - rye, P - pasture, and D.A. - diverted acres.

PER POINT AND PER FIELD CLASSIFICATION RESULTS

Data from Purdue Area C-3 July 1966

	Per Point Classifier			Per Field Classifier	
Class	No. of Samples	Percent Correct	60% Classifier	No. of Fields	Percent Correct
Soybeans	2380	60.6	60.5	13	84.6
Corn	1188	72.0	70.0	10	90.0
Mixture	2725	91.7	94.4	18	94.4
Water	101	97.0	100.0	3	100.0
	6394	80.3	81.2	44	92.3

Per Point Classifier: Overall Performance = 76.6

Per Field Classifier: Overall Performance = 90.9

Figure 24-18.- Per point and per field classification results for July, 1966 Purdue Flight Line C-3. The same training and test fields and the same spectral bands were used for both types of classification.

PER POINT AND PER FIELD CLASSIFICATION RESULTS

Data from Purdue Area C-4 July 1966

Class	Per Point Classifier: No. of Samples	Per Point Classifier: Percent Correct	Per Point Classifier: 60% Classifier	Per Field Classifier: No. of Fields	Per Field Classifier: Percent Correct
Soybeans	5302	83.0	89.5	19	89.5
Corn	8875	79.6	83.8	31	96.8
Pasture	5233	75.7	77.8	18	100.0
Stubble	2943	81.2	84.6	13	76.9
	22353	79.9	82.0	81	90.6

Overall Performance = 79.7

Overall Performance = 92.6

Figure 24-19.- Per point and per field classification results for July 1966 Purdue Flight Line C-4.

PER POINT AND PER FIELD CLASSIFICATION RESULTS

Data from Purdue Area C-2 September 1966

	Per Point Classifier			Per Field Classifier	
Class	No. of Samples	Percent Correct	60% Classifier	No. of Fields	Percent Correct
Soybeans	3804	51.4	66.7	12	66.7
Corn	3718	80.0	78.6	14	92.8
Pasture	3608	78.9	76.9	13	100.0
Stubble	3692	57.4	45.4	11	81.8
Water	134	99.3	100.0	3	100.0
	14956	73.4	73.5	53	88.3

Overall Performance = 67.0

Overall Performance = 86.8

Figure 24-20.- Per point and per field classification results for September 1966 Purdue Flight Line C-2.

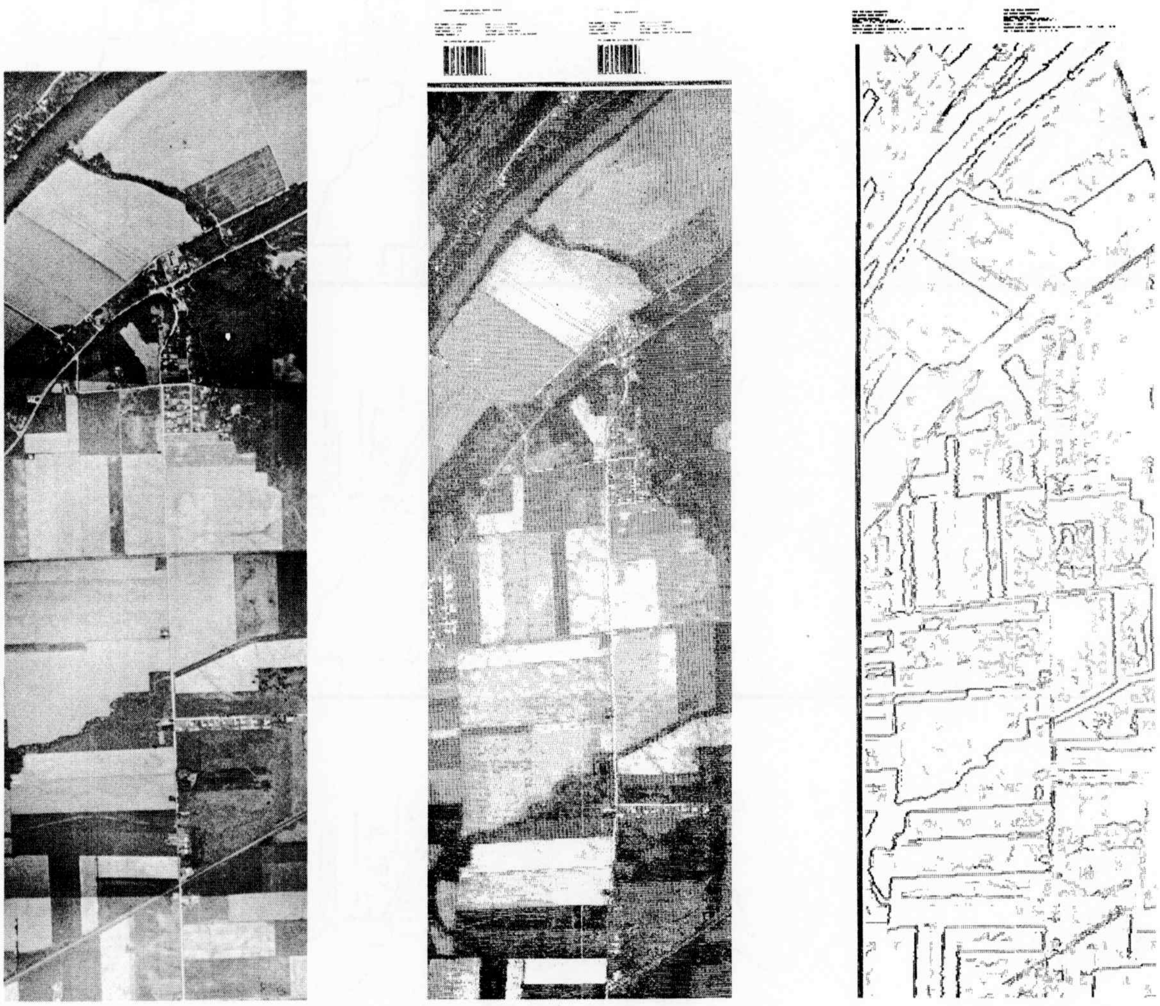


Figure 24-21.- Example output from the boundary-finding algorithm. An airphoto and line printer output are also given for comparison.

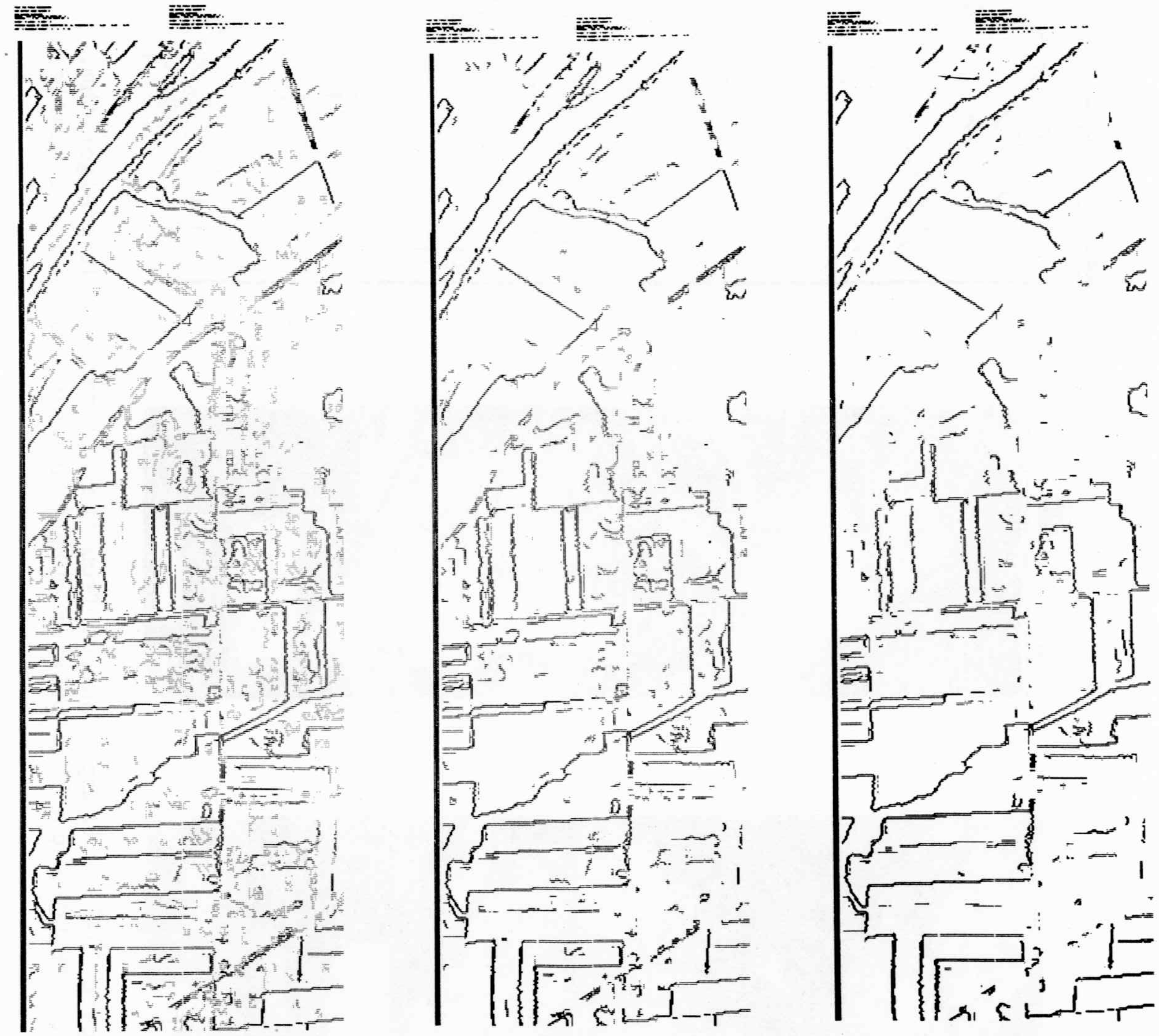


Figure 24-22.- Example results showing effects of varying the threshold of the boundary-finding algorithm.