Center for Advanced Computation

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN
URBANA, ILLINOIS 61801

124-24

CAC Document No. 168

SUMMARY OF ILLIAC IV - ARPA NETWORK MULTISPECTRAL IMAGE PROCESSING RESEARCH ACTIVITIES

Final Report, covering the period
March 1974 - February 1975

by

Robert M. Ray III

August 1975

CAC Document No. 168

SUMMARY OF ILLIAC IV - ARPA NETWORK MULTISPECTRAL IMAGE PROCESSING RESEARCH ACTIVITIES

Final Report,
covering the period March 1974 - February 1975
NASA Grant NGR 14-005-202

Robert M. Ray III
Principal Investigator
(217) 333-2228

of the

Center for Advanced Computation
University of Illinois at Urbana-Champaign
Urbana, Illinois 61801

August 1975

Reported research was conducted in collaboration with staff of the Laboratory for Applications of Remote Sensing, Purdue University, West Lafayette, Indiana 47906.

This work was supported by NASA's Ames Research Center in cooperation with the EROS Program of the U. S. Geological Survey, Department of Interior, through NASA Grant NGR 14-005-202 entitled: "Implementation of an ILLIAC IV Multispectral Image Analysis System". Collaborative support was also provided by the Statistical Reporting Service of the U. S. Department of Agriculture through USDA Research Agreement 12-18-04-8-1794-X, and by the Advanced Research Projects Agency of the Department of Defense through Contract DAHC04-72-C-0001.

ABSTRACT

This report summarizes ILLIAC IV - ARPA Network multispectral image processing research conducted during the last twelve months by the Center for Advanced Computation (CAC) of the University of Illinois at Urbana-Champaign. The research reported has focused on the implementation of ILLIAC IV - ARPA Network software systems for cost-efficient computer-assisted interpretation of multispectral earth-resources imagery such as that collected by NASA's LANDSAT satellites. This report overviews all multispectral image interpretation algorithms implemented to date specifically for ILLIAC IV processing of LANDSAT imagery and describes peripheral ARPA Network software systems developed to enable convenient, decentralized access to ILLIAC IV image analysis capabilities for a community of system users including NASA, USGS, and USDA.

CONTENTS

1. Introduction

This report summarizes multispectral image processing activities conducted during the period March 1974 through February 1975 by the Center for Advanced Computation (CAC) of the University of Illinois at Urbana-Champaign. This research has focused on the implementation of ILLIAC IV - ARPA Network software systems for economical computer-assisted interpretation of large quantities of earth-resources imagery such as that collected continuously by NASA's LANDSAT satellites. All of the research summarized within this report has been undertaken at CAC in collaboration with staff personnel of the Laboratory for Applications of Remote Sensing (LARS) of Purdue University, the Geographic Applications Program of the U. S. Geological Survey (USGS/GAP), NASA's Ames Research Center (NASA/Ames), and the Statistical Reporting Service of the U. S. Department of Agriculture (USDA/SRS). For the most part, support for this work has been provided by NASA's Ames Research Center, acting in cooperation with the EROS Program of the U. S. Geological Survey. Supplemental support has been provided by the Statistical Reporting Service of USDA.

The research activities summarized in this report represent a second-year continuation of CAC-LARS collaborative efforts initiated in March 1973 toward the design, partial implementation, and evaluation of ILLIAC IV - ARPA Network multispectral image processing systems for economical computer-assisted analysis of LANDSAT imagery.

During the first year of these efforts (March 1973 through February 1974), a subset of the computer procedures central to the LARS methodology of multispectral image processing[1,2] was selected for implementation on the ILLIAC IV. Specifically, LARS FORTRAN procedures for "unsupervised" interpretation of multispectral image data via multivariate cluster analysis techniques (see ref. 3.) and FORTRAN procedures for "supervised" classification of multispectral image samples via statistical decision-theory methods (again see ref. 3.) were redeveloped in parallel-computation mode and programmed in ASK for efficient execution on the ILLIAC IV.[4,5] Comparisons of the speeds and costs of the ILLIAC IV, the IBM 360/67, and the DEC PDP-10 computers with respect to these two multispectral image interpretation procedures indicated generally that the ILLIAC IV would be well over an order of magnitude more cost-effective than either the IBM 360/67 or the DEC PDP-10 for processing large quantities of LANDSAT imagery. The results of these initial demonstrations of the efficiencies of the ILLIAC IV for LANDSAT imagery analysis, as well as preliminary research findings concerning the potentials of the ARPA Network as a mechanism for decentralizing access to ILLIAC IV image processing capabilities for a nationally dispersed community of potential users, were documented both by CAC and LARS.[6,7]

Having demonstrated successfully the efficiencies of the ILLIAC IV for LANDSAT imagery analysis, and having evaluated favorably the potentials of the ARPA Network as a means of decentralizing access to ILLIAC IV LANDSAT data analysis capacities for a geographically dispersed community of government land-resources management agencies, support was provided to CAC for implementation of more comprehensive ILLIAC IV - ARPA Network image analysis systems.

The remaining sections of this report outline all progress made to date at CAC toward the development of these more complete ILLIAC IV - ARPA Network image analysis systems. In the next section, we overview briefly the EDITOR system -- a comprehensive DEC PDP-10 software system developed at CAC to provide a high-level-language, interactive ARPA Network interface to ILLIAC IV LANDSAT image interpretation procedures. In Section 3., we describe and document the efficiencies of all additional ILLIAC IV algorithms implemented at CAC for LANDSAT imagery analysis within the time period of this second-year project continuation. In Section 4., we describe all progress made to date toward the implementation of ARPA Network image data reformatting and management systems efficient for analysis, tabulation, and mapping of LANDSAT imagery. Finally in Section 5., we describe briefly the relationship of all past systems development work to present (third-year) CAC efforts toward the implementation of ILLIAC IV - ARPA Network software systems for cost-effective compilation and geographic registration of LANDSAT multitemporal imagery, and suggest areas of software development and refinement that we feel will require continued support in the future to achieve a fully-operational, comprehensive ILLIAC IV - ARPA Network LANDSAT image analysis facility.

2. EDITOR:
An Interactive Interface to ILLIAC IV - ARPA Network Image Analysis Systems

In this section, we present an overview of EDITOR -- ERTS Data Interpreter and TENEX Operations Recorder -- a comprehensive DEC PDP-10 (TENEX) software system now implemented to serve as a high-level-language, interactive ARPA Network interface to ILLIAC IV LANDSAT image processing capabilities.[8,9] Accessible via the ARPA Network, the EDITOR system allows portable, dial-up terminal command of ILLIAC IV image processing facilities at Ames. Multispectral image data management services within EDITOR have been designed to optimize the efficiency of ERTS data file transfers between the computers of the ILLIAC IV complex at Ames and other computers on the ARPA Network.

While the EDITOR system was conceived initially as a peripheral TENEX file-management and job-editing interface to ILLIAC IV batch image interpretations, the scope of EDITOR functions was subsequently widened to include procedures for direct TENEX analysis of small-scale LANDSAT image samples ("training" samples) in preparation for large-scale ILLIAC IV batch interpretations. Thus, where only small-scale image interpretations are needed and where portable terminal output is sufficient, the EDITOR system may be used alone as a self-sufficient tape-operating multispectral image analysis system. Specific image data tape formats required by EDITOR (as well as descriptions of all TENEX disk files generated by EDITOR) have been documented in CAC Technical Memorandum No. 19.[10]

The actual use of EDITOR is best described by way of examples. For this reason, complete records of three example EDITOR sessions illustrating all major system features have been included in CAC Document No. 114 -- the primary documentation to date of the EDITOR system.[9] The following verbal description of EDITOR simply sketches major system functions. (A familiarity with basic multispectral image interpretation procedures is assumed.)

EDITOR is implemented in modular format with each system module evoked by a single command. For any command, only that number of characters sufficient to distinguish the command from all others need be typed by the user.

The RETRIEVE command is used to create TENEX disk files containing multispectral image data from single rectangular areas of a LANDSAT image, or from sets of rectangular areas within an image. Use of RETRIEVE assumes that, prior to entering EDITOR, the user has requested the TENEX operator to mount a specific image tape. Analysis areas to be retrieved from tape are then specified by the user in terms of image line and column coordinates and line and column sampling increments.

If, on the other hand, the user has an image file already resident on the TENEX disk and wishes to use only some portion of it for analysis, he may use the CLIP A WINDOW command. This command allows the user to construct new windows by specifying new image line and column coordinates (and new line and column sampling increment as well) for the subwindows desired.

Once LANDSAT image analysis areas have been selected and stored on the TENEX disk, the PRINT command within EDITOR allows a user to view the image disk files created to verify that the desired data windows have indeed been retrieved. Gray-scale displays of any of the individual spectral bands of data within an image file may be achieved by overprinting characters on the user's terminal. A data histogramming routine is available for displaying the proportional distribution of pixel intensities within any spectral band of an image file. This feature of the system can be used for terminal display enhancement purposes. With reference to image data histograms, a user can select manually a particular range of image point intensities to be assigned to each gray-scale shade available within the character-overprinting terminal display facility.

Once a user has verified that the desired imagery has been selected, several options are available for data analysis.

The CLUSTER command within EDITOR evokes a multivariate cluster analysis of all multispectral image data points within a particular file to determine a logical partitioning or grouping of all data points into a user-specified small number of spectral clusters.[3,6] The cluster analysis procedures may be regarded as an unsupervised image interpretation procedure categorizing a set of image points into most-separable spectral classes.

Following any cluster analysis, a set of summary statistics characterizing the multivariate distributions of data within all clusters (i.e. mean vectors and variance-covariance matrices) is computed and, at the request of the user, displayed on the user's terminal.

The CLASSIFY command within EDITOR allows statistical classification of all data points within a specified disk file into a particular set of spectrally-distinct categories determined by a previous cluster analysis.

This methodology of classification assumes that "ground-truth" samples will always be cluster-analyzed prior to class signature computation in order to verify the existence of spectral homogeneity within nominal terrain classes and spectral separability between classes. Where spectral properties of ground-truth samples are not homogeneous within nominal classes (i.e. multivariate distributions are non-spherical), the multiple spectral classes determined by cluster analysis for such terrain classes may be used for analysis purposes as spectrally-distinct, but nominally-synonymous, classes to be re-grouped following classification into single nominal terrain classes.

In accordance with the LARS multispectral image interpretation methodology, the statistical classification algorithm evoked by the command CLASSIFY employs the Gaussian maximum likelihood decision rule.[1,2,3] For each data point of a file of data points to be classified, discriminant functions for all classes are computed and the point is assigned to that class for which the discriminant is largest. The parameters of the discriminant function for each class are the estimates of the means, variances and covariances of the spectral properties of all data points belonging to that class. (These are the spectral class statistics computed and saved by EDITOR immediately following any cluster analysis for the set of spectral classes determined by that particular CLUSTER procedure.)

The ENTER STATISTICS FILE EDITOR command provides a set of file-management procedures convenient for interactive display and editing of the spectral class statistics accumulated from previous cluster analyses. Within this mode, class signatures of different statistics files may be re-combined in any manner to produce new composite sets of signatures to be used for classification. Also within this mode, measures of spectral separability or distinctness may be computed and displayed for all pairs of signatures within any statistics file. A procedure is also implemented (again following LARS methods) for "pooling" the statistics of several different signatures into a single new composite signature.

The EXECUTE ANALYSIS ON ILLIAC IV command instructs EDITOR to use the ILLIAC IV at Ames for a particular cluster analysis or classification task specified by a user.[11] Given this command, EDITOR immediately requires the user's I4-TENEX user code and password. If such a code can be supplied, EDITOR will automatically set up the task for ILLIAC IV processing, transmit the job via the ARPA Network to Ames, and enter the job into the ILLIAC IV batch stream. Procedures are also provided within this mode for inquiring on the status of ILLIAC IV jobs and for convenient retrieval by EDITOR of cluster analysis and classification output files resulting from previous ILLIAC IV executions. Specific ILLIAC IV multispectral image cluster analysis and classification procedures accessible to EDITOR users have been documented elsewhere.[4,5,12,13]

In addition to the procedures for displaying raw multispectral imagery, the PRINT command within EDITOR also provides procedures for displaying the interpreted imagery that results from either cluster analysis or classification operations. Interpreted image data files may be output as terminal printer maps where the symbols 1 - 9 and A - Z are used to represent terrain classes. Display of only selected subsets of classes is permitted. Following again the LARS methodology, provisions are included for reliability thresholding of classification results within PRINT mode. Where a particular confidence level is specified by the user, EDITOR displays only those image data points whose classification reliability meets the specified level. Also, procedures are included within PRINT mode to allow re-grouping of spectral classes into composite nominal classes, and to allow manual re-assignment of printer symbols among these nominal classes for more meaningful display of interpretations.

3. ILLIAC IV Image Interpretation Algorithms

While the primary emphasis of the research conducted during the twelve-month period covered within this report has been placed on the development of interactive ARPA Network interfaces to ILLIAC IV image interpretation procedures and LANDSAT image reformatting and data management subsystems convenient for geographic referencing of LANDSAT imagery, additional ILLIAC IV algorithms have been developed as well to permit ILLIAC IV analysis of composite multitemporal LANDSAT image data sets.

The multispectral scanners (MSSs) aboard the LANDSAT series of earth-resources observational satellites collect digital imagery in the form of four-channel multispectral data sets. Since from the start we have attempted to implement all ILLIAC IV LANDSAT image interpretation procedures to achieve maximum processing efficiency, all algorithms developed during the previous twelve-month period of this research were implemented in ASK (the "low-level" machine language of the ILLIAC IV) specifically for interpretation of four-channel LANDSAT imagery. As pointed out above, the computational efficiencies of these procedures have been documented elsewhere.[6]

In the summer of 1974, however, experiments by LARS and USGS/GAP personnel indicated that, for purposes of regional land use inventories (via computer-assisted LANDSAT image interpretation methods), computer classification accuracies could be substantially increased by the use of multitemporal LANDSAT imagery, i.e., composite multispectral-multitemporal image data sets produced by "overlaying" one four-channel LANDSAT image taken over a particular region during one season with another four-channel image taken over the same region at some subsequent season. Since all requirements for digital processing tend to be substantially increased by this methodology, it was decided that additional ILLIAC IV procedures should be developed for analysis of eight-channel multitemporal imagery.

Thus, two additional ILLIAC IV subroutines were developed in the early fall of 1974: (1) an ILLIAC IV algorithm for cluster analysis of eight-channel LANDSAT imagery,[12] and (2) an ILLIAC IV algorithm for statistical classification of eight-channel LANDSAT imagery.[13] (Note: While general "N"-channel ILLIAC IV algorithms could be developed, due to the particular hardware structure of the ILLIAC IV and the specific format of all LANDSAT image data much computational efficiency is gained by standardizing all multitemporal-multispectral LANDSAT imagery to be processed on the ILLIAC IV in either four-channel or eight-channel format.)

Anticipating that it might become desirable at a later time to generalize further ILLIAC IV procedures for cluster analysis of LANDSAT data, it was decided to program the new eight-channel cluster analysis procedure in the higher-level language GLYPNIR rather than ASK. Since GLYPNIR permits only 64-bit arithmetic operations, however, this implied a substantial decrease in computational efficiency in comparison with the same procedure written in ASK which permits 32-bit arithmetic. Despite these considerations, it was decided to opt for the programming conveniences of GLYPNIR and hence, more possibilities for convenient modification of the program if subsequent alterations seemed desirable.

The new cluster analysis program[12] has been operational on the ILLIAC IV since October 1974. It was employed in the late fall of 1974 in support of USGS/GAP analyses of multitemporal LANDSAT imagery over the Washington, D.C. metropolitan area.

As an approximate bench mark of the efficiency of the ILLIAC IV GLYPNIR eight-channel cluster analysis procedure, 10,000 eight-channel LANDSAT image samples may be clustered into thirty-two (32) multispectral classes in approximately ten (10) minutes. (The algorithm is iterative and thus the exact time to convergence for a particular data set is somewhat unpredictable. Here we assume approximately 100 iterations for convergence -- a number close to the average for those cluster analyses of this scale that we have thus far performed.)

Given the relatively sound theoretical foundations of the Gaussian maximum likelihood decision rule underlying the statistical classification procedure, and since this procedure then represented the primary "workhorse" of the image interpretation methodology being implemented, it was decided to develop the additional eight-channel multitemporal classification algorithm again directly in ASK.[13] This new algorithm also became operational on the ILLIAC IV in the early fall of 1974.

In January of 1975, this algorithm was executed on the ILLIAC IV to classify into twenty-six (26) multispectral classes all image elements of an eight-channel multitemporal LANDSAT data set (compiled at LARS) for a region of approximately 4000 square miles surrounding Washington, D.C. This represented a first demonstration of the use of the ILLIAC IV system for interpretation of a full tape volume of multispectral imagery. The entire tape of 2,590,000 eight-channel LANDSAT image samples was classified on the ILLIAC IV in 478 seconds -- slightly less than eight minutes!

The statistical classification algorithm is non-iterative and hence the ILLIAC IV processing time for a specific data set can be predicted with considerable accuracy. In essence, the algorithm simply computes $\underline{m}$ discriminant functions ($\underline{m}$ is the number of spectral categories) for each of $\underline{n}$ image samples, and classifies each sample into that category for which the discriminant is largest. Thus, total ILLIAC IV time for any classification procedure is closely a linear function of $\underline{m} \times \underline{n}$, the number of discriminants to be computed. Extrapolating from the results of the Washington, D.C. classification, one may assume that discriminant functions will be computed for any eight-channel LANDSAT data at an effective rate of 7.1×10^{-6} seconds (7.1 microseconds) per discriminant. Note that this effective rate reflects as well the time required for ILLIAC IV disk data transfers.

4. Image Rectification and Data Management Systems

For a number of reasons, it is generally most convenient to work with the digital data of LANDSAT MSS imagery in terms of map-oriented rectangular image areas. For example, the image analysis area of interest is often the geographic region of a particular map bounded by lines of latitude and longitude. However, the characteristics of the near-polar orbit of the LANDSAT satellites are such that the grid of data elements scanned for each MSS image is always rotationally displaced to a considerable degree from alignment with geodetic meridians and parallels. Also, the rotation of the earth under LANDSATs during the scanning of each MSS image introduces an oblique skew to the grid geometry of each image digitized. Hence, to enable convenient retrieval, analysis, and display of selected MSS imagery in terms of map-oriented rectangular areas, cumbersome digital preprocessing is usually required to transform geometrically and reformat the computer tape data of each LANDSAT image to north-vertical image grid orientation.

This computer preprocessing of LANDSAT MSS data from initial scanner format to map-oriented image analysis format is known generally as image rectification or geometric correction. The two principle objectives of this data preprocessing step are (1) to remove the oblique skew within the MSS image grid caused by the earth's rotation during image scanning, and (2) to transform the MSS image grid to a vertical-north, map orientation. A third objective to be served by geometric-correction data preprocessing is sometimes the differential re-scaling of the MSS image grid along its transformed axes to achieve equality of horizontal and vertical image scales when the image is output using a particular display device.

Within the course of the research reported here, an efficient procedure for rectification of LANDSAT imagery has been developed based on the concept of image skew transformations.[14,15]

Skew transformation of an image can most easily be conceptualized as a vertical shifting of all image elements (pixels) of successive image columns by integral numbers of pixel positions so that resulting image rows best approximate linear sequences of data elements parallel to some specified vector, followed by a horizontal shifting of all new image rows so that, as well, successive pixel columns best approximate linear sequences perpendicular to row sequences. (For illustrations and further details see ref. 14. or 15.)

Such an image geometric-transformation procedure offers several distinct advantages for LANDSAT image rectification. First, where image displacements from desired axial alignments may be assumed to be small, the skew transformation provides directly for simultaneous image de-skewing and rotation to achieve map-oriented data formats. Second, since the mathematical mapping between input and output pixel coordinates is one-to-one, no data elements are lost (or replicated) by skew transformation. Furthermore, this mathematical mapping is inherently simple and always has an exact inverse. Third, digital implementations of the transformation on large (but generally available) scientific computers have proven to be surprisingly efficient.

Since the rotation of LANDSAT MSS imagery from north-vertical orientation is predictably small (approximately 13° over the U.S.), and since also image skew due to the earth's rotation is minor (on the order of 5%), the skew transformation offers a convenient strategy for rectifying entire LANDSAT images to map-oriented data formats.

In the fall of 1974, an IBM 360 PL/1 procedure for rectification of LANDSAT imagery via skew transformation was developed at CAC.[14] Since then, more than a dozen full-frame LANDSAT images have been geometrically corrected by CAC using this procedure for USGS, NASA, and USDA. This LANDSAT image rectification procedure is now performed routinely by CAC using either the IBM 360/91 computer at the University of California at Los Angeles via the ARPA Network, or locally, using the IBM 360/75 of the University of Illinois at a cost on the order of $100 per full LANDSAT image.

The PL/1 procedure for rectification of LANDSAT images via skew transformation requires as input two calibration numbers, namely the two slant parameters of image horizontal and vertical skew. These two parameters may be approximated directly as a function of available LANDSAT orbital information. More accurate estimates of these parameters, however, may be determined during establishment of more accurate geographic control for LANDSAT images by digitizing points of geographic control visually matched between LANDSAT film imagery and USGS maps.

The constancy of the relationship between LANDSAT digital image formats (2340 scanner lines with 3240 samples per line) and LANDSAT film image formats (photo-recordings of the center 2256 scan lines of corresponding digital images) facilitates greatly the use of film imagery for initial geographic calibration of digital imagery. Thus, by digitizing only a dozen or so control points identifiable on both the LANDSAT film image (at 1:500,000) and on USGS $1^{\circ} \times 2^{\circ}$ topographic maps (at 1:250,000), geographic control accurate to 5-10 pixels can be established simultaneously for film and digital image alike. Such an initial geographic calibration procedure requires typically well less than an hour of manual effort with readily available materials. Since such a procedure also allows more accurate determination of the horizontal and vertical skew parameters needed for skew transformation of an image to map orientation, it is usually convenient to calibrate the image-rectification data preprocessing procedure in this manner.

By way of recapitulation, geometric rectification of LANDSAT digital imagery via skew transformation offers the following distinct advantages for image analysis objectives:

1. it offers an efficient, low-cost procedure for geometric transformation of LANDSAT data to map-oriented formats that is executable on generally available scientific computers

2. it neither loses nor replicates individual image samples in transformation

3. it outputs a single image tape on which latitude-longitude quadrangles are ordered as rectangular blocks of data

4. line-printer displays of output data blocks corresponding to USGS 7 1/2' quadrangle maps at 1:24,000 overlay (fortuitously) these maps with close to one-pixel accuracy

5. a simple mathematical function relates original image coordinates to rectified image coordinates

Given that approximate geographic registration of imagery to map bases is provided directly by the manual procedure employed for calibrating the required skew transformation, and given that for small north-oriented rectangles of skew-transformed imagery line printer displays overlay USGS 7 1/2' quadrangles with essentially one-pixel accuracy, the problem of precision registration of LANDSAT imagery to USGS map bases becomes greatly simplified.

At this point, precision geographic control between digital imagery and map coordinates may be accomplished by selecting some set of areal features (river bends, lakes, highway intersections, etc.) visible within both the LANDSAT image to be precision registered and on available USGS 7 1/2' quadrangles, and by correlating visually small line-printer displays (say 64 lines by 64 columns) of the digital imagery with these same quadrangle areas to determine overlay positions of maximum geographic correspondence. A data tablet digitizer is then used to record these points of maximum geographic correspondence between digital image blocks and map quadrangles.

Experience to date suggests that approximately two dozen widely distributed points of geographic control established in the above manner are sufficient to achieve one-pixel geographic registration accuracy over an entire LANDSAT image. Once such a set of control points has been established, a precision transformation between image line and column coordinates and latitude and longitude geo-coordinates is computed and stored as a TENEX file associated with the LANDSAT image thus precision registered. This precision transformation file is then available to any user working subsequently with the LANDSAT image and, via the CALCULATE COORDINATES command within EDITOR, any user is able to geo-reference any particular area of the LANDSAT image to one-pixel accuracy through specification (or digitization) of the area in terms of map coordinates.

5. Present and Future System Developments

In the preceding sections of this report, we have summarized progress made roughly within the twelve-month time period March 1974 through February 1975 toward implementation of ILLIAC IV - ARPA Network LANDSAT image analysis systems. In this concluding section, we overview briefly research activities currently under way within present third-year system development efforts, offer a tentative evaluation of all systems to be completed by the end of this present research period, and suggest several areas that we feel will deserve continued NASA/USGS support in the future to attain fully the operational capabilities of ILLIAC IV - ARPA Network LANDSAT image analysis systems developed.

Current CAC image processing research activities supported by NASA, USGS, and USDA are concentrated generally within three major areas: (1) the development of more economical computer-assisted methods for compilation and geographic registration of LANDSAT multitemporal data sets, (2) the development of numerical methods for efficient geometric transformation of LANDSAT digital data to image formats conforming to USGS map projections, and (3) the development of ARPA Network TENEX software procedures (as extensions of the EDITOR system) to support on-line digitizing and interactive data management of the substantial amounts of ground-truth information typically associated with LANDSAT imagery analysis.

Previously (see ref. 15.) we have reported on software developments presently under way at CAC toward the implementation of comprehensive ARPA Network - ILLIAC IV facilities for compilation and geographic registration of multitemporal LANDSAT imagery.

The composite system under development takes full advantage of the opportunities for computational efficiency afforded by a heterogeneous computer network such as the ARPA Network. The ILLIAC IV may be employed for calculation of the numerous image-to-image block correlations computations required for accurate digital overlay of temporally separate images as multitemporal data sets. The large IBM 360/91 at UCLA is utilized efficiently for basic image rectification operations (skew transformations), as well as for all image data tape reformatting required. Interactive ARPA Network TENEX systems (as described above) are used for on-line computational support of all image-to-map overlay digitizing tasks necessary for accurate geographic registration of LANDSAT digital imagery.

The throughput of the system as presently developed at CAC is somewhat limited in that it relies completely on computational facilities remote and dispersed with respect to the image analyst and hence, requires mail service transmission of digital image tapes between component processor sites. On the other hand, the system achieves considerable economic efficiency by judicious decomposition of all processing required into a sequence of individual tasks well-suited to a particular combination of processors currently available via the ARPA Network. In a manner completely satisfactory for research purposes (i.e., one-week turnaround), the system makes possible the overlay and geographic registration of seasonal pairs of LANDSAT images for computer costs on the order of $.5K per full image pair (or $.6K where IBM 360/91 processing is substituted for ILLIAC IV processing in calculating image-to-image block correlations.) At the time of this writing, the system has been used successfully in compiling multitemporal LANDSAT data sets of seasonal pairs of images over the Seattle and San Francisco Bay regions.

Closely allied with this image registration research is an investigation of the feasibility of a comprehensive LANDSAT image registration facility to be based at NASA/Ames. To the extent that the recently acquired CDC 7600 at Ames proves efficient for the image transformation and reformatting tasks now done on the IBM 360/91, it should be possible to implement on computers resident at NASA/Ames essentially the same LANDSAT image registration system developed at CAC for ARPA Network "distributed" computing.

Toward a general goal of producing economically LANDSAT imagery analysis products of maximum benefit to land resources management agencies, research continues at CAC toward the derivation, programming, and evaluation of numerical procedures for efficient cartographic transformation of LANDSAT digital imagery (both raw and interpreted) to image formats conforming to USGS map projections.

As pointed out above, bulk rectification of LANDSAT imagery to vertical-north orientation via the skew transformation method results in image tapes from which USGS latitude-longitude quadrangles of imagery may be retrieved conveniently as (approximately) rectangular blocks of data. Following precision geographic registration of this rectified imagery, digital image files conforming more precisely in geometry to specific USGS maps may be created by reformatting further the rectified imagery corresponding to the ground area of a particular map in accordance with the numerically determined transformation relating image element coordinates to latitude-longitude coordinates and a further mathematical transformation specific to the cartographic projection of the particular map being fitted.

The major problem of research interest here is to develop numerical transformations for specific map projections that represent appropriate compromises between the conflicting objectives of computational efficiency and map projection accuracy. To date, "compromise" transformations have been derived for both the polyconic and transverse mercator (UTM) projections used by USGS for the national topographic series of 7 1/2', 15', and 1° x 2° map quadrangles. A CAC technical memorandum describing this work is forthcoming.[16]

Using as primary ground-truth source materials USDA/SRS low-altitude aerial photos over agricultural regions of Kansas and Texas, work is proceeding toward the development of additional EDITOR subroutines to support on-line digitizing and interactive data management of ground-truth information recorded on aerials and USGS quadrangle maps.[17] Systems now in experimental use support on-line digitizing of polygonal ground-truth fields (both convex and non-convex) and interactive identification, retrieval, and aggregation of all LANDSAT image samples falling within particular subsets of fields specified by the user in terms of a hierarchy of nominal and geographic field identifiers. While all systems currently operational for interactive ground-truth digitizing and management reflect strongly the specific needs of USDA/SRS for agricultural crop-acreage monitoring, more generalized versions of these same procedures should prove advantageous to EDITOR system users within NASA and USGS as well.

In addition to our efforts to implement ILLIAC IV algorithms for LANDSAT imagery interpretation following methods previously researched at LARS, we have also explored alternative mathematical bases for pictorial pattern information processing. Aside from the purely theoretically interesting nature of the problem, our primary motivation here has been the general conviction (shared by almost every researcher in the area) that substantial increases in computer classification accuracies currently obtained with traditional statistical techniques for interpretation of earth-resources reconnaissance imagery will come about only as new methodologies are developed that employ the spatial (contextual) as well as spectral information content within the imagery to be interpreted.

Within the general context of our research on parallel-computation algorithms for image processing, we have come across a particular mathematical model of pattern information processing that works quite well for certain pattern recognition applications where previous methods have performed rather poorly.[18] This particular mathematical approach to spatial pattern information processing we are now researching lends itself well to parallel computation, and we hope to investigate in the near future applications of the general model to several specific pattern information processing tasks intrinsic to the general problem of computer interpretation of multispectral earth-resources imagery.

The majority of the LANDSAT image analysis systems that we have discussed within this report may already be considered operational, productive, and economically efficient. EDITOR is now used routinely, not only by CAC, but by personnel of USGS, NASA and USDA as well. As early as last January within the context of the USGS/GAP Washington, D.C. regional land-cover inventory and mapping experiment, the efficiencies of the ILLIAC IV for full-frame multitemporal LANDSAT imagery interpretation were convincingly demonstrated and documented.[19] Though still being refined, the basic capabilities and efficiencies of the ARPA Network - ILLIAC IV systems for multitemporal LANDSAT imagery compilation and geographic registration were demonstrated in May 1975 for imagery over the Seattle region. We have consistently sought to modify and refine all software systems developed, as practical, in accordance with feedback from NASA, USGS, and USDA system users and have provided freely copies of any software procedure requested for operation elsewhere.

While we feel much has already been accomplished toward achievement of operational ILLIAC IV - ARPA Network LANDSAT image analysis systems (and we expect many additional accomplishments during the next six months), we anticipate that continued FY-76 support of CAC image processing activities will be warranted within perhaps several of five general areas:

1. continued maintenance, refinement, and further development of all LANDSAT image analysis systems previously developed
2. assistance to NASA/Ames in establishing comprehensive "on-base" image registration systems and providing service image registration support during the interim period
3. work toward development of information management systems specifically designed for the potentials of regional change detection and urban growth monitoring operations made practical by ILLIAC IV image interpretation capabilities
4. research ILLIAC IV implementations of multispectral image interpretation methods developed by LARS, the University of Kansas, and CAC based on spatial (textual, contextual) as well as spectral information processing
5. continued refinement of computer graphics techniques for operational production of interpreted LANDSAT imagery overlays for USGS topographic maps and other base map series used for regional planning purposes

REFERENCES

1. K. S. Fu, D. A. Landgrebe, and T. L. Phillips, "Information Processing of Remotely Sensed Agricultural Data", Proceedings IEEE, Vol. 57, No. 4 (April 1969), pp. 639-653.

2. "Remote Multispectral Sensing in Agriculture", Research Bulletin 873, (December 1970), Laboratory for Agricultural Remote Sensing, Purdue University, West Lafayette, Indiana 47906.

3. P. H. Swain, "Pattern Recognition: A Basis for Remote Sensing Data Analysis", LARS Information Note 11572, (1972), Laboratory for Applications of Remote Sensing, Purdue University, West Lafayette, Indiana 47906.

4. J. Thomas, "An ILLIAC IV Algorithm for Cluster Analysis of ERTS-1 Data", CAC Technical Memorandum No. 17, (May 1974), Center for Advanced Computation, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801.

5. J. Thomas, "An ILLIAC IV Algorithm for Statistical Classification of ERTS-1 Imagery", CAC Technical Memorandum No. 18, (May 1974), Center for Advanced Computation, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801.

6. Robert M. Ray III, John D. Thomas, Walter E. Donovan, and Philip H. Swain, "Implementation of ILLIAC IV Algorithms for Multispectral Image Interpretation", CAC Document No. 112, (June 1974), Center for Advanced Computation, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801.

7. Philip H. Swain, "Implementation and Evaluation of ILLIAC IV Algorithms for Multispectral Image Processing", LARS Technical Memorandum T-16 111273, (July 1974), Laboratory for Applications of Remote Sensing, Purdue University, West Lafayette, Indiana 47906.

8. Robert M. Ray III, "An Interim Description of EDITOR -- An Interactive Interface to ILLIAC IV ERTS Image Analysis Systems", CAC Technical Memorandum No. 55, (July 1974), Center for Advanced Computation, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801.

9. Robert M. Ray III, Martin Ozga, Walter E. Donovan, John D. Thomas, and Marvin L. Graham, "EDITOR: An Interactive Interface to ILLIAC IV - ARPA Network Multispectral Image Processing Systems", CAC Document No. 114, (June 1975), Center for Advanced Computation, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801.

10. Martin Ozga, Walter E. Donovan, Robert M. Ray III, John D. Thomas, and Marvin L. Graham, "Data File Formats for Processing of Multispectral Image Data", CAC Technical Memorandum No. 19, (July 1975), Center for Advanced Computation, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801.

11. John Thomas and Martin Ozga, "RUN-I4: Controlling Execution of ILLIAC IV ERTS Analysis Programs from a Remote TENEX Site", CAC Technical Memorandum No. 34, (September 1974), Center for Advanced Computation, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801.

12. Ozga, Martin, "An ILLIAC IV Algorithm for Cluster Analysis of 8-Channel Multispectral Data", CAC Technical Memorandum No. 53, (June 1975), Center for Advanced Computation, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801.

13. Ozga, Martin, "An ILLIAC IV Algorithm for Statistical Classification of 8-Channel Multispectral Data", CAC Technical Memorandum No. 54, (June 1975), Center for Advanced Computation, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801.

14. Walt Donovan, "Oblique Transformation of ERTS Images to Approximate North-South Orientation", CAC Technical Memorandum No. 38, (November 1974), Center for Advanced Computation, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801.

15. Walter E. Donovan, Martin Ozga, and Robert M. Ray III, "Compilation and Geographic Registration of ERTS Multitemporal Imagery", CAC Technical Memorandum No. 52, (May 1975), Center for Advanced Computation, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801.

16. Walter E. Donovan, "Numerical Procedures for Reformatting LANDSAT Images to Overlay USGS Maps of Specific Cartographic Projections", CAC Technical Memorandum No. 56, (August 1975), Center for Advanced Computation, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801 (in preparation).

17. Walt Donovan and Martin Ozga, "Retrieval of LANDSAT Image Samples by Digitized Polygonal Windows and Associated Ground Data Information", CAC Technical Memorandum No. 57, (August 1975), Center for Advanced Computation, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801.

18. Robert M. Ray III, "On Methods for Direct Quantification of Pattern Associations", CAC Document No. 139, (November 1974), Center for Advanced Computation, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801.

19. Leonard Gaydos, "Method of Completing ERTS Computer-Aided Land-Cover Analysis of the Washington, D. C. Metro Area", USGS/GAP Internal Memorandum, (February 1975), Geographic Applications Program, U. S. Geological Survey, Menlo Park, California 94025 (mimeo).

REPORTS AND PUBLICATIONS

(Accumulation to date for NASA Grant NGR 14-005-202)

1. R. Ellefsen, P. Swain, and J. Wray, "Urban Land-Use Mapping by Machine Processing of ERTS-1 Multispectral Data", Proceedings of Conference on Machine Processing of Remotely Sensed Data, (October 1973), Purdue University, Lafayette, Indiana.

2. J. Thomas, "An ILLIAC IV Algorithm for Cluster Analysis of ERTS-1 Data", CAC Technical Memorandum No. 17, (May 1974), Center for Advanced Computation, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801.

3. J. Thomas, "An ILLIAC IV Algorithm for Statistical Classification of ERTS-1 Imagery", CAC Technical Memorandum No. 18, (May 1974), Center for Advanced Computation, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801.

4. R. Ray and J. Thomas, "ILLIAC IV Execution Times for Two ERTS-1 Data Interpretation Algorithms", CAC Technical Memorandum No. 13, (September 1973), Center for Advanced Computation, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801.

5. Robert M. Ray III, John D. Thomas, Walter E. Donovan, and Philip H. Swain, "Implementation of ILLIAC IV Algorithms for Multispectral Image Interpretation", CAC Document No. 112, (June 1974), Center for Advanced Computation, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801.

6. Philip H. Swain, "Implementation and Evaluation of ILLIAC IV Algorithms for Multispectral Image Processing", LARS Technical Memorandum T-16 111273, (July 1974), Laboratory for Applications of Remote Sensing, Purdue University, West Lafayette, Indiana 47906.

7. Robert M. Ray III, "An Interim Description of EDITOR -- An Interactive Interface to ILLIAC IV ERTS Image Analysis Systems", CAC Technical Memorandum No. 55, (July 1974), Center for Advanced Computation, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801.

8. John Thomas and Martin Ozga, "RUN-I4: Controlling Execution of ILLIAC IV ERTS Analysis Programs from a Remote TENEX Site", CAC Technical Memorandum No. 34, (September 1974), Center for Advanced Computation, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801.

9. Walt Donovan, "Oblique Transformation of ERTS Images to Approximate North-South Orientation", CAC Technical Memorandum No. 38, (November 1974), Center for Advanced Computation, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801.

10. Robert M. Ray III, "On Methods for Direct Quantification of Pattern Associations", CAC Document No. 139, (November 1974), Center for Advanced Computation, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801. (accepted for publication in Pattern Recognition, the Journal of the Pattern Recognition Society)

11. Walter E. Donovan, Martin Ozga, and Robert M. Ray III, "Compilation and Geographic Registration of ERTS Multitemporal Imagery", CAC Technical Memorandum No. 52, (May 1975), Center for Advanced Computation, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801.

12. Martin Ozga, "An ILLIAC IV Algorithm for Cluster Analysis of 8-Channel Multispectral Data", CAC Technical Memorandum No. 53, (June 1975), Center for Advanced Computation, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801.

13. Martin Ozga, "An ILLIAC IV Algorithm for Statistical Classification of 8-Channel Multispectral Data", CAC Technical Memorandum No. 54, (June 1975), Center for Advanced Computation, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801.

14. Robert M. Ray III, Martin Ozga, Walter E. Donovan, John D. Thomas, and Marvin L. Graham, "EDITOR: An Interactive Interface to ILLIAC IV - ARPA Network Multispectral Image Processing Systems", CAC Document No. 114, (June 1975), Center for Advanced Computation, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801.

15. Martin Ozga, Walter E. Donovan, Robert M. Ray III, John D. Thomas, and Marvin L. Graham, "Data File Formats for Processing of Multispectral Image Data", CAC Technical Memorandum No. 19, (July 1975), Center for Advanced Computation, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801.

16. Walt Donovan and Martin Ozga, "Retrieval of LANDSAT Image Samples by Digitized Polygonal Windows and Associated Ground Data Information", CAC Technical Memorandum No. 57, (August 1975), Center for Advanced Computation, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801.