10

Special
E7.5 - 1 0.3 6 2
CR-143320

Interim Report

ORSER-SSEL Technical Report 4-75

PROCESSING ERTS AND AIRCRAFT MSS DATA WITH THE GENERAL ELECTRIC IMAGE 100 SYSTEM

B. F. Merembeck and F. Y. Borden

ERTS Investigation 082
Contract Number NAS 5-23133

INTERDISCIPLINARY APPLICATION AND INTERPRETATION OF ERTS DATA WITHIN THE SUSQUEHANNA RIVER BASIN

Resource Inventory, Land Use, and Pollution

Office for Remote Sensing of Earth Resources (ORSER)
Space Science and Engineering Laboratory (SSEL)
Room 219 Electrical Engineering West Building
The Pennsylvania State University
University Park, Pennsylvania 16802

(E75-10362) PROCESSING ERTS AND AIRCRAFT MSS DATA WITH THE GENERAL ELECTRIC IMAGE 100 SYSTEM Interim Report (Pennsylvania State Univ.) 8 p HC $3.25 CSCL 05B

N75-29504

Unclas
G3/43 00362

Principal Investigators:

Dr. George J. McMurtry
Dr. Gary W. Petersen

1082A

Date: June 1975

PROCESSING ERTS AND AIRCRAFT MSS DATA WITH THE GENERAL ELECTRIC IMAGE 100 SYSTEM

B. F. Merembeck and F. Y. Borden

In accordance with the provisions of NASA contract number NAS 5-23133, ORSER undertook an evaluation of the General Electric Image 100 System. Briefly, the Image 100 is an interactive multispectral image display and analysis system incorporating a color cathode ray tube (CRT) with a fast refresh disc in combination with a minicomputer. The minicomputer interfaces the CRT disc with the multispectral scanner (MSS) data tapes, with the system software held on a program disc. Classification results are output on up to eight theme tracks which are stored on, and recalled from, the refresh disc for display.

Objectives

The initial intent of the investigation was two fold: first, to use standard ERTS-MSS tapes as received from NASA to observe the performance of the system using data for which it was designed; and second, to use canonically transformed tapes generated from aircraft and ERTS data, in a format compatible with the Image 100 system. The objective was to observe the general performance of a color CRT with a variety of data sources, and to determine if the transformed tapes would improve the accuracy and speed of classification.

Date Transformation and Preliminary Analysis

With these objectives in mind, it was clear that most of the preparation by ORSER would involve developing and implementing the software to generate nonstandard data tapes in ERTS format. The Image 100 system will accept only ERTS format tapes of four channels at 800 BPI. Non-ERTS MSS tapes are currently available in a number of formats, such as MSS-DAS[1], LARSYS II[2], etc. The program library at ORSER contains a number of programs which use data from the above source tapes. In the interest of efficiency, it is desirable that these programs operate with a common data format and with smaller subsets of the original data tapes.

To meet these requirements, the original SUBSET program was developed. This program generates working subset tapes in ORSER format from the original source data tapes. As applied to aircraft MSS data, this facility has processed Houston C130 Bendix scanner data, in both LARSYS II and MSS-DAS formats, and Reconofax IV thermal scanner data. For the Image 100 project, SUBSET was extended

[1]MSS-DAS is the format system developed by NASA at the Johnson Space Flight Center, Houston, Texas.

[2]LARSYS II is the format system used by the Laboratory for Agricultural Remote Sensing.

to reformat data tapes compatible with the SUBSET program into ERTS format tapes. The actual format order in the program is SOURCE DATA - ORSER - ERTS. The result is a four-channel tape with 804 elements per scan line in ERTS format. Details of this process are presented in ORSER-SSEL Technical Report 16-74: ERTS and Aircraft Multispectral Digital Data Users Manual.

The initial Image 100 session for ORSER personnel was held on February 9-10, 1974. This was mainly for familiarization with the system. Primary concentration was focused on a portion of an ERTS scene of the East Branch Reservoir of the Clarion River in northwestern Pennsylvania (scene 1028-15295, 20 Aug 72). As an interface between the interpreter and his data, the color CRT was clearly superior to the computer line maps ORSER has been using. Classification of this area using the Image 100 1-D classifier was visually judged to be reasonably accurate by an interpreter familiar with the scene. However, a second attempt, by the same interpreter, to classify the scene could not duplicate the results of the first classification. The problem here was the necessity of relying on the visual cursor to input training coordinates.

The only nonstandard ERTS tape brought on this trip was one with recalibrated data from the East Branch scene. This was used to determine whether the system would accept ERTS format tapes generated by the SUBSET program. The system accepted the tape and some classification was performed. Limited time on the Image 100 prohibited a complete analysis.

In preparation for subsequent work on the Image 100, the SUBSET program was further extended to incorporate the following procedures:

1. Averaging

   With only four channels available in the ERTS format and up to 24 available from the Bendix scanner, the Bendix channels were averaged to approximate the channel responses of ERTS. This was both for comparison and for reducing the number of data channels to conform to ERTS format.

2. Data Transformation

   The transformation is of the form $A\underline{x} = \underline{Z}$ where A is an M row by N column matrix, $\underline{x}$ is the raw MSS element of N channels, and $\underline{Z}$ is the transformed element of M axes. Hence A is a linear transformation which converts an N channel raw MSS element into an M axes transformed element. This feature is specifically intended for use with the canonical transformation matrix generated by the CANAL program in the ORSER library, but any conformable matrix may be used. The canonical transformation increases the interclass distances between elements of a given set of classes, and often reduces the intraclass distances as well.

ERTS and Aircraft Data Analysis

The next session on the Image 100 (June 12-13, 1974) was again devoted mainly to original ERTS tapes. ERTS scenes from the Shamokin (scene 1350-15190, 8 Jul 73) and Lancaster (scene 1080-15185, 11 Oct 73) areas of Pennsylvania and a scene from Texas (scene 1146-16323, 16 Dec 72) were analyzed with both the 1-D and N-D classifiers. In addition, the Texas scene was read in several times with different radiometric scale factors to determine the effect of the scale factors on the channel means and variances for the same training area. It appeared that rather than being a linear effect (as assumed by GE), a change in the radiometric scale factor changed the variance of the data by the square of the scale factor. This could be predicted statistically.

The only non-ERTS tape contained aircraft data scanned at 5000 ft. from the McElhattan area on the West Branch of the Susquehanna River (NASA C130 Mission 238, Flightline 4). The source of this tape was 12 channels of data from the Bendix scanner, in MSS-DAS format. The 12 channels of the source data were averaged into four channels of ERTS format data with a response range from 0.38 micrometers to 2.3 micrometers. The system accepted this tape. The spatial resolution was excellent, approaching photographic quality while simultaneously providing a data base in digital form. Photographs were taken of the face of the CRT and a scratch tape left with General Electric.

The third session took place on July 1-2, 1974. This session was devoted exclusively to aircraft data. The areas and data configurations were as follows:

1. McElhattan area - West Branch Susquehanna River - approximately 5000 ft. - 12 channels. (NASA C130 Mission 238, Flightline 4)

   a. Twelve channels of data in MSS-DAS format averaged into four channels in ERTS format.

   b. Twelve channels of data in MSS-DAS format canonically transformed onto four axes and output in ERTS format.

2. Pine Creek area - West Branch Susquehanna River - approximately 15,000 ft. - 8 channels. (NASA C130 Mission 230, Flightline 5)

   a. Four channels of data in MSS-DAS format output as four channels in ERTS format.

   b. Six channels of data in MSS-DAS format canonically transformed onto four axes and output in ERTS format.

3. North Bend area - West Branch Susquehanna River - approximately 5000 ft. - 14 channels. (NASA C130 Mission 230, Flightline 16)

   a. Twelve channels of data in MSS-DAS format averaged into four channels in ERTS format.

   b. Fourteen channels of data in MSS-DAS format transformed by the four eigenvectors associated with the four eigenvalues accounting for 98.2 percent of the variability in the scene and output on four axes in ERTS format.

   c. Fourteen channels canonically transformed onto four axes and output in ERTS format.

This was the first attempt at detailed analysis using aircraft data. The high spatial resolution allowed use of the displayed image alone to select training areas. No supplementary aircraft photography was needed. The 1-D classification was performed on both the averaged and transformed data. The classification on the transformed data was faster and, in the case of non-homogeneous targets such as forests, more consistent than classification of nontransformed data. At the resolution afforded by the aircraft data, each pixel in a forest scene has an area roughly equal to one tree crown or less at small look-angles of the scanning mirror. This results in very nonhomogeneous response characteristics for forest data. Classification of nontransformed data both by ORSER classification methods and on the Image 100 was erratic. However, 1-D classification of the transformed data gave very good and consistent results when applied to the McElhattan scene.

Density slicing of the first canonical axis gave good classification of the areas of interest in both the McElhattan and North Bend scenes. In the Pine Creek scene the second axis gave better delineation of the floodplain areas of interest. This is perhaps due to some incorrect training areas specified in preparing statistical information to be input to the canonical analysis program. The incorrect training areas were of a forested section next to the river interpreted as floodplain. This conceivably caused the first axis to separate forest land from nonforest land and water.

This success at uniform density slicing indicated that even better results could be achieved if the slicing levels for specific categories were initially determined by ORSER and then input to the Image 100 system. This was put on the schedule for the final session.

The N-D classifier was also run on the aircraft data. The high spatial resolution of the data resulted in alarms (categories) with orders of magnitude of 10^5 cells. With this number of cells in a single alarm, the search procedure of the N-D classifier becomes quite time consuming. For this reason N-D classification of the larger alarms was discontinued.

The final session on the Image 100 system took place on September 26, 1974. Emphasis was placed on the analysis of transformed data. Both ERTS and aircraft data were processed, as follows:

1. North Bend area - West Branch Susquehanna River - aircraft - 5000 ft. - 14 channels (NASA C130 Mission 230, Flightline 16)

    a. Twelve channels of MSS-DAS data averaged into four channels in ERTS format.

    b. Fourteen channels of MSS-DAS data canonically transformed onto four axes and output in ERTS format.

2. Whipple Dam State Park - Central Pennsylvania - ERTS - eight channels (ERTS scenes 1459-15223, 25 Oct 73, and 1297-15252, 16 May 73).

    a. Data from the two scenes were merged onto an eight channel tape in ORSER format. The eight channels were then canonically transformed onto four axes and output in ERTS format.

3. Gypsy Moth Infestation - Northeastern Pennsylvania - ERTS - four channels (ERTS scene 1350-15183, 8 Jul 73).

   a. The four channels of data were transformed onto two canonical axes. The first axis was repeated three times to output a four channel tape in ERTS format.

4. West Branch Susquehanna River - aircraft thermal - 5000 ft. - 1 channel (NASA C130 Mission 238, Flightlines 1&4).

   a. The one channel was repeated three times, and outpu , along with the calibration channel, in ERTS format.

The desired levels for density slicing the first axis of the transformed data were initially determined by ORSER. The results were quite satisfactory, particularly for the gypsy moth scene. Although signatures for moderately defoliated forest had not been developed to input to the canonical analysis program, it was known from earlier studies that the moderately defoliated signature lay between the healthy and heavily defoliated signatures on untransformed data. A density slice of the first canonical axis between the healthy and heavily defoliated intervals of the axis successfully mapped the moderately defoliated areas.

As a further system evaluation, the cluster analysis algorithm of the Image 100 system was run on the averaged North Bend data. As with the N-D classifier, the high spatial resolution of the aircraft data resulted in large cell numbers in the alarm and the clustering was discontinued, as prohibitively time-consuming, after 14 iterations.

Evaluation

Observations and suggestions based on ORSER's experience with the Image 100 system as of September, 1974, are discussed here. It should be borne in mind that the system was in a state of continual development during the evaluation.

Advantage of the Color CRT as an Interface Between the Interpreter and His Data

The CRT is less abstract than other display systems. Compared to a computer line map, for example, mental juggling is not required to interpret the characters of the line map. This facilitates location and definition of training areas.

The utility of the CRT can be enhanced in two ways. First, a method of determining local spectral uniformity would be useful. This would allow faster and more consistent training area delineations, and should reduce most of the time-consuming and expensive histogram slicing required with the I-D classifier.

Second, the interpreter should have the option of directly addressing cell areas by line and element coordinates. It should be possible to input training area corners via the Tektronix keyboard, and to quickly recall line and element coordinates of previous training areas. Without this, results are not readily repeatable.[3] It is too cumbersome to read coordinates off the Tektronix after visually setting a training area with the joystick. At $250/hr, any hinderance to fast user access of data in the system, or to fast input of data to the system, must be considered a deficiency.

Strengths and Weaknesses of Using a Dedicated Minicomputer on Bulk MSS Data

The dedicated computer allows a very interactive system. Combined with the visual definition of the color CRT, it allows a rapid sequence of interpreter decisions. This advantage deteriorates, however, as processing becomes more complex. Time constraints are imposed by the frequent paging to and from storage on the refresh disc, necessary because of the limited storage capacity of the minicomputer. This was a particularly significant deficiency when processing aircraft data. Large alarms resulted in lengthy processing during search procedures of the N-D and cluster analysis programs. During such processing, the system is not available for interaction with the interpreter.

It would be desirable to run the more sophisticated and complex programs on a large general purpose machine, where the processing time would be only a fraction of that required with the minicomputer. Results could be output directly to the system or to tape which could be read into the system. Use of a large computer would minimize the amount of bulk processing required of the minicomputer and increase the overall efficiency and capability of the Image 100 system. The interpreter could then have access to the interactive facilities of the Image 100 while the bulk processing was being done externally. Charges for the IBM 370/168 used by ORSER are $366/cpu and system hour; hence the $250/external clock hour for the Image 100 system further favors the larger machine for bulk processing.

Finally, the Image 100 system would be more effective if documentation were available to the interpreter prior to use of the machine. As mentioned before, the system was in a continuous state of development during this investigation. As no documentation was available during this period, ORSER was never fully cognizant of the capabilities of the system at any given time. This caused considerable difficulty, as one had to learn how to use the system from the results obtained by trial and error. It is strongly recommended that system documentation be made available to future users.

[3] This deficiency caused considerable problems in comparing the Image 100 display with ORSER computer output. The ORSER signatures and the Image 100 themes were not developed from the same training areas, hence an element by element comparison was impossible.

7 Special

ABSTRACT

ORSER-SSEL Technical Report 4-75

PROCESSING ERTS AND AIRCRAFT DATA WITH THE GENERAL ELECTRIC IMAGE 100 SYSTEM
B. F. Merembeck and F. Y. Borden

The General Electric Image 100 is an interactive multispectral image display and analysis system incorporating a color cathode ray tube (CRT) with a fast refresh disc in combination with a minicomputer. The minicomputer interfaces the CRT disc with the multispectral scanner (MSS) data tapes, holding the system software on a program disc. Classification results are output on up to eight theme tracts which are stored on, and recalled from, the refresh disc for display.

The objectives of this project were two fold: 1) to observe the performance of the system using the data for which it was designed -- that is, ERTS(Landsat)-1 data tapes; 2) to determine if the processing of aircraft MSS data, and aircraft and ERTS MSS data which have been linearly transformed, could improve the speed and accuracy of classification of the Image 100 System. The System will accept only ERTS format tapes of four channels at 800 BPI.

Using the ORSER data processing system at Penn State, selected aircraft data from the 24-channel Bendix scanner were both averaged to approximate the channel responses of the four ERTS channels and linearly transformed by a process which converts a raw MSS element of N channels into a transformed element of M axes by means of a canonical transformation matrix. The same transformation process was performed on selected ERTS data sets.

The Image 100 accepted the various data tapes prepared by ORSER. The spatial resolution was excellent, approaching photographic quality while providing a data base in digital form. The classification on transformed aircraft data was faster, and in some cases more consistent, than classification of non-transformed data. Classification of nontransformed aircraft data gave generally erratic results. Uniform density slicing of the transformed data for specific categories gave excellent results. These categories included healthy, moderately defoliated, and heavily defoliated forest areas effected by gypsy moth infestation; and a floodplain area.

The color CRT is an excellent interface between the interpreter and his data. It is less abstract than other display systems, facilitiating location and definition of training areas. The utility of the system could be improved by 1) the interpreter having the option of directly addressing cell areas by line and element coordinates, without which results are not readily repeatable; and 2) a method of determining local spectral uniformity permitting faster and more consistent training area delineations. The dedicated computer combined with the visual definition of the color CRT permits a rapid sequence of interpreter decisions. However, the limited storage capacity of the minicomputer is a definite disadvantage, particularly when processing aircraft data. Use of a large general purpose computer would cut processing time and cost to a fraction of that required with the minicomputer and the interpreter could have access to the interactive facilities of the Image 100 while bulk processing was being done externally.