CR-142203

ENVIRONMENTAL RESEARCH INSTITUTE OF MICHIGAN

FORMERLY WILLOW RUN LABORATORIES, THE UNIVERSITY OF MICHIGAN

P. O. BOX 618 • ANN ARBOR • MICHIGAN • 48107 PHONE (313) 483-0500

101900-48-L
24 February 1975

Developing Processing Techniques for Skylab Data
Monthly Progress Report, January 1975

EREP Investigation 456 M
NASA Contract NAS9-13280

Prepared by

Richard F. Nalepka - Principal Investigator

William A. Malila - Co-Principal Investigator

James P. Morgenstern

NASA Technical Monitor

Mr. Larry B. York/TF6
National Aeronautics and Space Administration
Johnson Space Center
Principal Investigator Management Office
Houston, Texas 77058

(E75-10153) DEVELOPING PROCESSING TECHNIQUES FOR SKYLAB DATA Monthly Progress Report, Jan. 1975 (Environmental Research Inst. of Michigan) 5 p HC $3.25 CSCL 14E

N75-18665

Unclas
G3/43 00153

PRECEDING PAGE BLANK NOT FILMED

Developing Processing Techniques for Skylab Data
Monthly Progress Report, January 1975

The following report serves as the twenty-third monthly progress report for EREP Investigation 456 M which is entitled "Developing Processing Techniques for Skylab Data". The financial report for this contract (NAS9-13280) is being submitted under separate cover.

The purpose of this investigation is to test information extraction techniques for SKYLAB S-192 data and compare with results obtained in applying these techniques to LANDSAT and aircraft scanner data.

Twin processing efforts continued during January on the SKYLAB S-192 multispectral scanner data and the aircraft M-7 multispectral scanner data.

PROCESSING OF SKYLAB DATA

The month's work on SKYLAB data continued the S-192 data quality analysis reported last month and began an effort to locate specific fields and areas in the S-192 data for use as training and test fields.

As regards the former, nothing was found to change the preliminary conclusions regarding signal-to-noise problems reported last month. In response to our request for conic-scan data, we promptly received a data tape which, unfortunately, did not include the requested area. We re-requested the correct area in conic data format.

We proceeded to the job of locating fields and other geographical features. Graymaps of several of the bands displayed good contrast and homogeneous areas were clearly evident. However, upon close inspection, it was not possible to accurately find many of the boundaries between fields. Additionally, it was not possible to discern other geographic features, e.g., roads, so that we could not accurately match our ground information with the graymap.

It was therefore decided to locate specific fields and other points of interest in the scene by means of a procedure where all points of interest (section corners, field corners, etc.) are located on large scale photography and the (x,y) coordinates of these points are calculated. These (x,y) coordinates are then transformed to data (scan line, scan point) coordinates. A procedure [1] to do this was developed at ERIM for use in LANDSAT data processing, where similar field definition problems had occurred.

[1] Malila, W. A., R. H. Hieber, and A. P. McCleer, "Correlation of ERTS MSS Data and Earth Coordinate Systems", Proc. of Purdue Conf. on Machine Processing of Remotely Sensed Data, Oct. 1973, (NTIS No. N74-13037).

To carry out the (x,y) measurements in a fast and precise manner we used the Bendix Data Grid system which efficiently digitizes coordinates of points where a cursor has been momentarily positioned.

After some study, it was decided to use large-scale black-and-white enlargements of color-IR imagery acquired by the U-2 overflights in mid August, 1973. Two frames will have to be used to cover the whole of the test area.

Obviously, to employ such a system, it is necessary to provide a mechanism to transform photographic (x,y) coordinates into data (scan line, scan point) coordinates. This is done by using control points, i.e., specific points in the photograph which also can be found accurately in the graymaps. Being unable to find in the graymaps such sources of control points as roads or intersections, it was decided to use bodies of water as sources of control points. A comparison of signatures for a deep water lake and a general vegetative area showed a large separation of signals in two bands, SDO 17 (1.15-1.28 μm) and SDO 19 (.93-1.05 μm) for these two classes. We proceeded to produce a likelihood map for water (i.e., printing a map where the symbol indicates the probability of the data point being water) based on these two bands. In this manner we were able to locate precisely both lakes and muck fields (very wet soil) in the data. Using S-190A color-IR and U-2 color-IR imagery we were able to locate the corresponding lakes and muck farms in the enlarged U-2 photographs.

To this time, we have completed the spotting of all points of interest in one of the U-2 photos. We began by indicating all section corners for 40 sections and then carefully delineated the boundaries of all fields within a section which were large enough to be visible in S-192 data. An average of 11 fields per section were marked in this manner. Finally, specific control points from the water areas were carefully chosen and marked on the photos. In the coming month we intend to digitize all the points on the one photo and complete the location of fields in that area before we begin working on the second photo.

PROCESSING OF AIRCRAFT DATA

We continued training procedures on the aircraft-acquired data set. In the end, we had three separate sets of training signatures. What follows is an explanation of each set, how and why it was acquired, and an assessment of its utility. In all of what follows, we are talking only about corn, tree and soybean signatures, although it should be understood that we continued to use signatures for other classes (hay, weeds, soils, etc.) as described in last month's report. Also, all training and testing described below was performed in the same area, which was outlined in last month's report.

First is the set of signatures described last month. These are signatures formed by combining signatures of cluster groups of common object classes. We had combined two clusters for corn, and three each for trees and soybeans. The result of performing classification with these signatures on the training area showed a number of tree classifications in corn fields. Also, overall, the correct classification rate was only fair.

It was thought that part of the problem, especially the corn-tree problem, might have come from combining the cluster groups into one signature per class. Possibly one of the tree clusters represents some sparse tree area which would have spectral characteristics very similar to corn. Thus, another classification was carried out using one signature from each cluster. Overall correct classification decreased a bit from the previous set; additionally, it was found that the tree-corn confusion was limited to two of the three tree clusters. The third cluster signature recognized trees almost exclusively and accounted for half of the tree points in the scene. Of the points classified as the first cluster signature, 40% were from tree points and 30% from corn points. For the second of the three so-called tree clusters, almost 60% of the points classified to it were from corn fields and only 30% from trees. How, then, did these two clusters become associated with trees? A second look at the cluster map showed that while some of the points in these two clusters were from corn fields, the majority were from tree areas. Moreover, most of the corn points which were later classified to one of these two tree clusters were originally assigned to one of the two corn clusters. We believe that what has happened is a demonstration of tracking phenomena -- i.e., in clustering the means of the clusters are constantly changing, being influenced by the newer additions to the cluster. Thus it may be that the corn clusters changed enough during the clustering of the data so that some points which were once associated with one cluster may be classified as belonging to a second cluster. In any event, this set of signatures resulted in slightly poorer recognition accuracy and slightly increased false alarm rates over the first discussed set of signatures above.

As a final investigation we wanted to contrast the results obtained using cluster-derived signatures against some obtained using the "classical" training set approach. So we extracted signatures from all the corn, soybean and tree areas in the training area. All the signatures for each class were combined, after first omitting "outlying" signatures, i.e., signatures whose mean was further than some distance θ from the mean of the combined signatures. It was found that signatures thus discarded were from anomalous fields -- tree areas which were pasture with some trees, a soybean field that was very "ratty", etc. Results using this set showed a marked increase in correct classification. The tree-in-corn misclassification was much smaller than for the other tests. Also, where several misclassifications occurred in otherwise homogeneous areas, it was found from examining photography that these matched up with ditches, dead spots in fields, or other actual inhomogeneities in these areas. Thus, we found that the best set of classification signatures for the aircraft data is this last set.

We plan to classify the entire data set during the coming month using the usual linear decision rule. We may also want to perform classification using some of the newer classification schemes developed at ERIM such as the adaptive classifier, or one of the nine point classification rules.

Submitted by: Richard F. Nalepka
Richard F. Nalepka
Principal Investigator

Approved by: Jon D. Erickson
Jon D. Erickson
Head, Information Systems
and Analysis Department

Approved by: Paul Rice for
Richard R. Legault,
Director, Infrared and
Optics Division

dd