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

101900-52-L
21 April 1975

Developing Processing Techniques for Skylab Data
Monthly Progress Report, March 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

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

The following report serves as the twenty-fifth 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.

The SKYLAB S-192 data set being studied under this contract is the same data set being studied by a group from Michigan State University under the direction of Dr. Lester V. Manderscheid. Inasmuch as ERIM has subcontracted with MSU to perform S-192 data analysis under contract NAS9-13332, a monthly report similar to this reporting the common data preparation is also being issued under that contract.

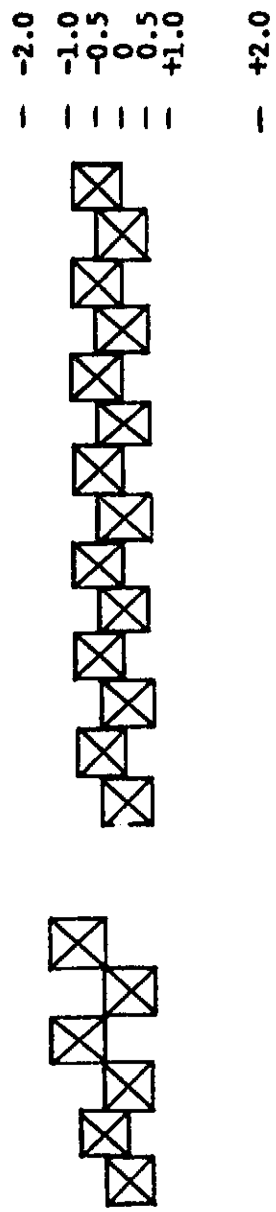
During the reporting period we continued to emphasize the data preparation aspects of the task in an effort to finish entirely this part of the work. The decision to perform most of the processing on the non-scan line straightened, or conic, format data (which is discussed below) caused this process to be drawn out further than had been anticipated. Jobs performed during the month included extending field location to conic scan line and point coordinates, the marking and digitization of field location points from the second U-2 acquired imagery, investigations into SDO to SDO misregistration in the conic and straightened data and measurement of pixel size for the S-192 data for the Michigan EREP test site data.

To begin with, we were concerned over the effects of misregistration of SDOs on processing of S-192 data. A report [1] issued concerning S-192 sensor evaluation called out 4 SDOs which were not perfectly registered. Examining the conic data, we found a sizeable water body where three scan lines made the transition from land to water at the same point. Signals from these scan lines were averaged and the resulting data normalized to the boundary value and plotted. The results, shown in Figure 1, showed the same misregistration problems as reported in the reference. We attempted a similar analysis on the straightened data, however we could not locate lake areas where the boundary occurred at the same point on several consecutive scan lines. An analysis was made using just individual scan lines

FIGURE 1. MISREGISTRATION OF CONIC FORMAT S-192 DATA

(The symbols (☒) indicate the relative projection on the ground for all SDO's for one resolution element.)

SDO	Misregistration (±.25 Resel)
1*	-0.5
2*	0
3	-0.5
4	0
5	-0.5
6	0
7	-0.5
8	0
9	-0.5
10	0
11	-0.5
12	0
13	-0.5
14	0
15	Not Used
16	Not Used
17	-1.0
18*	0
19	-1.0
20	0
21*	-0.5
22*	0



*Indicates Bands where the land water boundary is indistinguishable; reported is the assumed correct registration characteristic.

but this failed because the noise in the data made the individual scan line traces too irregular to analyze. We could reach no conclusion regarding the misregistration of the straightened data.

It was therefore decided to pursue a more analytical technique to measure the misregistration of the data. An algorithm was developed which calculates the cross correlation function for a pair of SDOs for many different simulated degrees of misregistration between the SDOs. The actual misregistration between the SDOs is that for which the cross correlation function is a maximum. The algorithm is detailed in Appendix 1 to this report. We plan during the coming month to implement the algorithm by programming it for our computer, and then using the program to measure the misregistration for both the conic and straightened data sets.

However, it was felt that further processing should not be suspended pending the outcome of the above program. Since we wanted to continue with the processing effort and since we were convinced of the existence of misregistration, it was decided to continue the processing effort on the conic data. We feel that we could correct the conic data for misregistration since the algorithm to correct for misregistration is simple for the conic data and we felt that we had a good estimate of the misregistration of the conic data. We would further mention here that we have experienced very little problem in working with the conic data especially in regard to obtaining line and point numbers of particular pixels. Since the test area is located near the middle of the data swath, we have found that ordinary graymaps are only slightly distorted and are eminently useable for our needs.

Having decided to process the conic data it was next necessary to convert the previously obtained straightened data line and point numbers for the fields in the ground information area to conic line and point numbers. This was done by using the inverse of the scan line straightening transformation equations as given in the EREP Users Handbook, coupled with regression techniques to accurately calculate the constants in the equations. The equations we used are:

$$\text{CONIC POINT} = A\left[\frac{N}{\theta}\sin^{-1}\left[\frac{P\pi\theta}{180\cdot N}\right] + \left[\frac{N+2}{2}\right]\right] + B$$

where

P = [STRAIGHT POINT - 517.8-0.5]

N = 1239 Points/Conic Scan Line

θ = 116.25° Field of Scan

A & B are constants estimated from regression techniques.

Similarly, for scan lines:

$$\text{CONIC LINE} = C + D \cdot \text{STRAIGHT LINE} - E \cdot R \cos\left[\frac{(\text{CONIC POINT} * 2 - 2 - N)\theta}{2N}\right]$$

with

R = Radius of the scan circle projected on the Earth

R $\approx$ 608 pixels

and C, D, and E are constants estimated from regression techniques. To perform the regression, 18 points were located on both conic and straightened graymaps. The regression fit was very good and further, all 5 coefficients seemed to be sensible, a reflection of the physical reality.

With the field coordinates converted, the ground information was merged with the conic data. Graymaps of two conic data channels and the ground information channels were overlayed for comparison and the conversion was deemed very satisfactory.

Finally, we ran the data through a deskewing program to reduce the misregistration in the data. With reference to Figure 1, we took the even numbered, high sample rate SDOs, along with all the low sample rate SDOs. Thus there were three channels out of registration with the rest, two of these by a full pixel. The registration algorithm shifted these two channels over by one full pixel. The other SDO was registered by estimating the signal expected midway between the two samples. Initially the estimator used was a simple average of the two adjacent samples.

During the reporting period we completed the marking of fields and other points of interest on the second U-2 acquired photograph, and digitized all these points as we had done for the first photograph. Again following the procedure established while working with the first photo, regression analysis was performed to calculate parameters to convert from photo coordinates to scan line straightened coordinates. Then the digitized points were converted. The next step is to convert, again using regression techniques to estimate parameters, to conic line and point numbers and to then merge the ground information with the data. We plan to do this during the next month.

Other work during the period included the following three items:

1) We received screening film of all SDO's (except for 15 and 16 which are redundant) and examined it thoroughly. We expect this to be a useful tool throughout the analysis of the data.

2) A brief analysis of actual pixel (not reselm) size was conducted. Pairs of pixels in lakes were located on straightened data graymaps which were either on the same scan line, several hundred pixels apart, or located at the same scan point number several hundred scan lines apart. Points corresponding to the pixels were also located on USGS maps of Southern Michigan. Distances were accurately measured on the USGS maps and on the graymaps with the result that the pixels were measured to be 69 meters wide along scan direction and 72 meters in the along track direction. Calculations based on geometrical considerations using only the angle of the scan cone and the altitude at the time of data acquisition yielded measures of 70 x 70 meters. The EREP Users Handbook calls out the pixel size to be 72 x 72 meters. The differences are not felt to be serious.

3) Crop acreage was totaled for each of the 90 sections in the ground information area and grand totals were also calculated. This showed the predominant ground covers and percentage of total area as:

Corn	30%
Trees	17%
Grasses	25%
Stubbles	9%
Soil	7%

During the next reporting period, in addition to those items already mentioned above, we intend to extract spectral signatures and perform analyses of the signatures, including calculation of optimum bands. With a set of signatures we will then be able to begin work on applying the mixtures classifier to SKYLAB data. We will also be able to evaluate signature extension techniques as applied to SKYLAB data. Further, we also intend to study the effects of misregistration on classification results for scan line straightened data where misregistration may not be correctable.

Reference:

[1] "ERIM Contributions to the S-192 Sensor Performance Evaluation", John G. Braithwaite and Peter F. Lambeck, ERIM 102800-51-F, January 1975.

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

APPENDIX 1

The following is a procedure for determining the amount of misregistration between two correlated data channels. By reconstructing the continuous waveform over a lengthy interval in both channels, the cross-correlation function of the two waveforms can be determined. Let f(t) and g(t) denote the reconstructed waveforms in the two channels over the interval [A,C]. The cross-correlation function $r(t_o)$ is defined as

$$r(t_o) = \int_A^C f(t)g(t + t_o)\, dt$$

The amount of misregistration between the two channels can be estimated as the value of the parameter t_o which maximizes the cross-correlation. The continuous waveforms can be reconstructed from the sample values by making assumptions which allow the use of Shannon's sampling theorem. The sampled data is converted into continuous form to allow the misregistration to be estimated to within a fraction of a pixel rather than in whole pixel increments. The length of the interval [A,C] must be long in comparison to the range of the parameter values t_o. This condition is required to minimize the effect of inaccuracies which will occur near the endpoints of the interval.

Shannon's sampling theorem indicates that a continuous signal y(t), bandlimited to B(radians/sec), can be exactly reconstructed from samples taken with a sampling interval $\tau = \pi/B$. The sampling rate is equal to twice the highest frequency component contained in the signal. The original signal y(t) can be expressed in terms of the sample values $y(m\tau)$ as

$$y(t) = \sum_{m=-\infty}^{\infty} B\ y(m\tau)\ \frac{\sin B(t - m\tau)}{B(t - m\tau)}$$

Assume that the two continuous data channels f(t) and g(t) are bandlimited to B and that the sampling interval τ is equal to π/B. Let the sample values of these two waveforms over the interval [A,C] be denoted as $f(k\tau)$ and $g(i\tau)$ $i, k = 1, \ldots, N$. The cross-correlation $r(t_o)$ can be expressed in terms of the samples as

$$r(t_o) = \sum_{i,k=1}^{N} f(k\tau)g(i\tau) \int_A^C \frac{\sin B(t - k\tau)}{B(t - k\tau)} \frac{\sin B(t + t_o - i\tau)}{B(t + t_o - i\tau)} dt$$

Using a variation of Parseval's Theorem, the integral can be evaluated by extending the limits of integration to positive and negative infinity, and $r(t_o)$ can be expressed as

$$r(t_o) = B\pi \sum_{i,k=1}^{N} f(k\tau)g(i\tau) \frac{\sin B(k\tau - i\tau + t_o)}{B(k\tau - i\tau + t_o)}$$

or, since $B\tau = \pi$

$$r(t_o) = B\pi \sum_{i,k=1}^{N} f(k\tau)g(i\tau) \frac{\sin \pi(k - i + t_o/\tau)}{\pi(k - i + t_o/\tau)}$$

This relationship can be expressed in terms of a fraction of a sampling interval (or fraction of a pixel) by defining a variable $\Delta = t_o/\tau$. Then

$$r(\Delta) = B\pi \sum_{i,k=1}^{N} f(k\tau)g(i\tau) \frac{\sin \pi(k - i + \Delta)}{\pi(k - i + \Delta)}$$

Neglecting the constant factor $B\pi$ and expressing $f(k\tau)$ and $g(i\tau)$ as f_k and g_i, respectively, the function $d(\Delta)$ must be evaluated, where

$$d(\Delta) = \sum_{i,k=1}^{N} f_k\, g_i \frac{\sin \pi(k - i + \Delta)}{\pi(k - i + \Delta)}$$

which can be simplified as

$$d(\Delta) = \sum_{j=-(N-1)}^{(N-1)} \left[\sum_{i-k=j} f_k \, g_i \right] \frac{\sin \pi(\Delta - j)}{\pi(\Delta - j)}$$

For large N, the variable j need not extend over the entire range because of the insignificant contribution of the high magnitude terms. To reduce the effects of noise, the function $d(\Delta)$ should be determined for several scan lines and averaged.