1-1-1981

Image Registration System in the Landsat-D Production Environment

Peter Kiss

Paul Arnold

John Goldstine

Reprinted from

Seventh International Symposium

Machine Processing of

Remotely Sensed Data

with special emphasis on

Range, Forest and Wetlands Assessment

June 23 - 26, 1981

Proceedings

Purdue University
The Laboratory for Applications of Remote Sensing
West Lafayette, Indiana 47907 USA

IMAGE REGISTRATION SYSTEM IN THE LANDSAT-D PRODUCTION ENVIRONMENT

PETER KISS, PAUL ARNOLD, JOHN GOLDSTINE

General Electric Company
Lanham, Maryland

This work was supported by NASA under contract no. NAS5-25300

I. SUMMARY

An effort is underway to develop an automated image registration system for the Landsat-D Ground Segment. This system will be capable of providing accurate control point (CP) location errors in imagery that has been corrected using system models. A part of this effort consisted of studying various image enhancement techniques, correlation techniques and subpixel registration methods. Presented here is an overview of the registration system developed, along with the study results that led to the choices of techniques incorporated.

Although much previous work exists in this area, it is believed that some of the methods and findings are new. It is also hoped that the extensive testing results along with the constraints of a very high speed production environment will be of value to the remote sensing community.

II. SYSTEM OVERVIEW

The flow chart in figure 1 gives a global view of the registration process. Below we give a brief description of most of the steps involved.

The purpose of the system is to take pieces of imagery, called control point chips (CPC), whose geodetic location has been previously determined and stored, and locate their position in later imagery of the same area. The registration processes are carried out partially on a DEC VAX 780 computer and partially on a Floating Point Systems Array Processor (AP-120B). Typically sets of 20 control points are processed at a time. To process these as sets, and to optimize the use of both machines, operations are grouped into loops instead of a sequential processing for each point.

After initializations and preparations of parameters ①, the first set of operations ② consists of: 1.) radiometric correction of the raw imagery; 2.) a pseudo cloud cover assessment; and 3.) geometric correction, in the horizontal direction, using system models. This loop and the next two are done in the AP120.

The next two loops, ③ & ④, which are essentially identical, are the heart of the registration. For non-cloud covered imagery, we do the following: 1.) Use a gradient operator to enhance both the CP chip and CPN; 2.) Form a correlation surface with the two using a normalized cross correlation similarity measure; 3.) Sort the correlation surface according to height; and 4.) Compute the mean and standard deviation of those heights.

The difference between loops ③ and ④ is the imagery size, called the control point neighborhood (CPN), in which the chip is located. After three consistent correlations are achieved with the full sized CP neighborhoods, we reduce their size and thus reduce computation time to about ¼ of the original for the remaining control points.

In the last loop ⑤ we do: 1.) Clustering to determine whether secondary peaks are present; 2.) Subpixel location determination; and 3.) The generation of some quality data and houskeeping data. There are also various tests along the way to determine pixel and subpixel registration success and to test for outliers.

III. CLOUD COVER ASSESSMENT

In order to save on computation time and to reduce the number of false correlations, a pseudo cloud cover assessment technique was developed. It is called pseudo because it is incapable of differentiating clouds from other sources which saturate pixels of the imaging sensor, such as snow. However, it is hoped that the technique will be able to differentiate clouds and/or snow from other bright sources which are constant in an area, such as sand.

When a control point chip is originated for library storage, it is extracted by an operator interactive system, and is checked visually to be cloud free. At this time, a histogram of the area is created, and the percentage of pixels above a certain brightness threshold is computed. This number gives us the quantity of normally present saturation in the neighborhood of the control point.

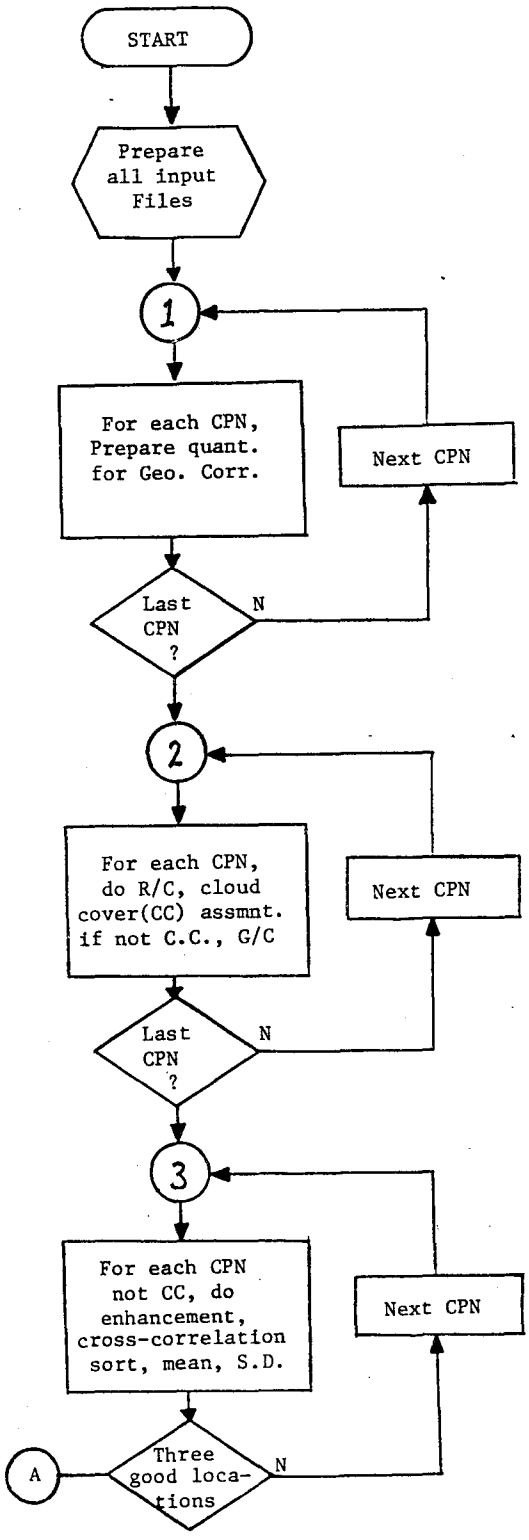

Figure 1a. FLOW CHART OVERVIEW

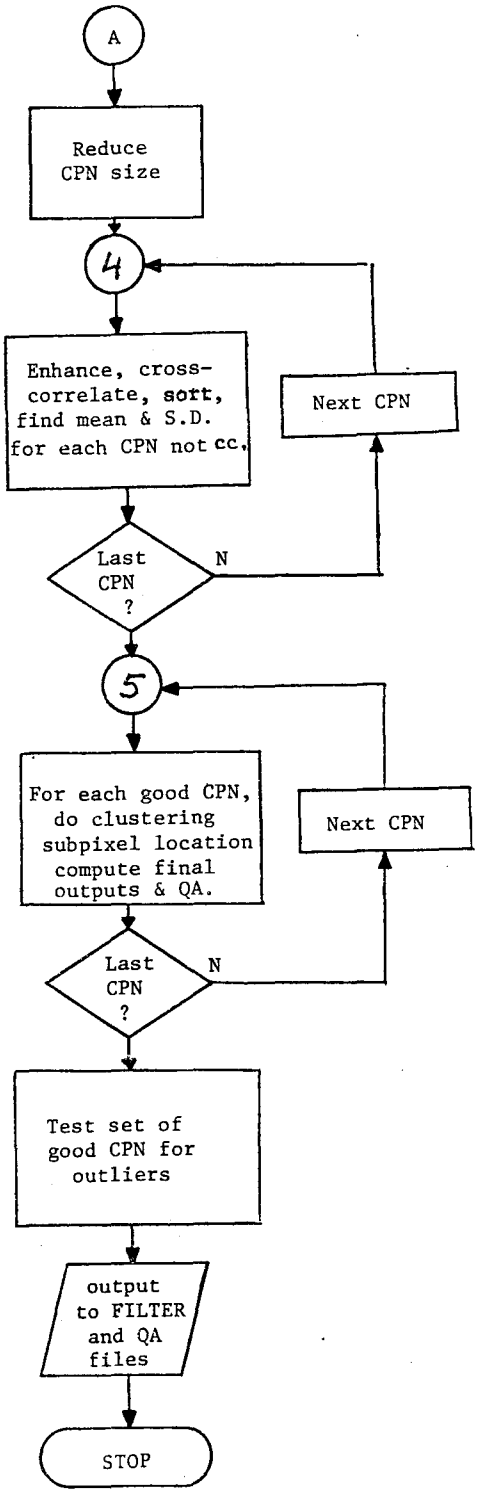

Figure 1b. FLOW CHART OVERVIEW

When a later piece of imagery is to be registered to this control point chip, the same threshold, T, is used to compute the percentage of saturated pixels in it. The two percentages are compared, and if the later image has a sufficiently higher saturation level than the original, it is rejected. Thus no further processing is done on images which are deemed cloud (or possibly snow) covered.

IV. ENHANCEMENT

The basic ingredients of the enhancements consisted of: gradient and Laplacian masks as edge detectors; a histogram equalization technique to sharpen images; and thresholding criteria to form edge extraction.

For a local array of radiance values

$$\begin{matrix} r_1 & r_2 & r_3 \\ r_4 & r_5 & r_6 \\ r_7 & r_8 & r_9 \end{matrix}$$

the gradient mask used is

$$g(r_5)=(|r_1-r_9|+|r_2-r_8|+|r_3-r_7|+|r_4-r_6|)/4 \quad (1)$$

At each location (such as r_5) the neighboring pixels (r_2, r_4, r_6 and r_8) are checked against upper and lower thresholds, t_U and t_L, which have cloud and shadow characteristics. Only if all the tests are negative is the gradient computed. The purpose of these checks is to avoid adding the edges caused by clouds and shadows to the gradient masked images, since they would greatly degrade the correlation.

Several other gradient type operators were tried, including

$$g_1(r_5)=\max\{|r_1-r_9|,|r_2-r_8|,|r_3-r_7|,|r_4-r_6|\} \quad (2)$$

and

$$g_2(r_5)=\left[(r_2-r_8)^2+(r_4-r_6)^2\right]^{\frac{1}{2}} \quad (3)$$

In all cases a non-directional mask was desired so as not to bias the gradient image. Table 1 shows results of a few correlations using the three gradients. The g_1 technique was considerably worse than the other two, while g (the one picked) was comparable to the true gradient, g_2, at a saving of computation time.

The Laplacian mask used is

$$L(r_5)=r_2+r_4+r_6+r_8-4r_5 \quad (4)$$

Again this was found about as good as others while being economical computationally.

For histogram equatization, the original image histogram is generated, and the original radiance range (0-127) is divided into N (say N=20) intervals.

Table 1.

Registration Strength of Gradients
(Maximum = 1, Acceptable $\geq$ .5)

Image Set	g_1	g	g_2
1	.823	.864	.888
2	-1.059	.221	.173
3	.859	.898	.908
4	.881	.920	.853
5	-6.022	-1.361	-2.631
6	.691	.808	.820
7	-3.926	-2.157	-1.075

Then histogram values are distributed into these N bins in such a way that each bin has the same number of pixels in it, thus equalizing the histogram. New radiance values are assigned according to the bin a pixel ended up in. The result is a distribution of contrast in high contrast areas and a lumping together in low ones.

Two modes of thresholding were implemented, a deterministic mode and a statistical mode. In the deterministic mode, a percentage was specified and the program sets all pixels with radiance values in the top x (say 15) percent to one and all others to zero. In the statistical mode, the values of one and zero were assigned according to distance (in units of standard deviation) from the mean, i.e. pixels whose radiance values are greater than 1 σ from the mean are set to one, all others to zero.

The above techniques were used in combinations to produce more complex enhancement methods. The gradient followed by a thresholding constitutes an edge extraction technique, resulting

in a binary image with ones outlining the edges of our original image. By varying the percent thresholded, we can control the amount of edges in the enhanced image. The smaller the amount of edges in an image the sharper the correlation will be. But the risk of non-registration due to such effects as rotation, increases accordingly. Nominally, about 15% of an image might be considered edges. The thresholding technique was also tried after the use of a Laplacian. It is not clear what to call this combination, physically it extracts non-linear changes (or edges).

Other sets of enhancements were created by spplying histogram equalization before and after using the gradient or Laplacian. This was also tried in combination with the thresholding.

Since this system is being designed for a production mode, our data set was chosen to have a great deal of variety. Three bands of Multispectral Scanner (MSS) imagery were explored. Different feature types were used as control points. Seasonal variations were allowed. Because of these variables, and the rather harsh acceptance criteria (explained later), the ratio of registration is not very high, but the results are believed to be realistic.

In Table 2 we summarize the results of the various enhancements in use with the normalized cross-correlation similarity measure. Notice that the combination of histogram equalization with gradient and Laplacian do not yield any measurable improvements, and in many cases degrade results. Also the addition of histogram equalization lowers registration success when used with thresholding of any form. It appears that the Laplacian is favored in band 5, the gradient in band 6, and both are good in band 7. Thresholding the gradient reduces its power by a small amount at the gain of computational ease on most computers (not the AP120). Thresholding the Laplacian resulted in very strange results (not sorted out here), which should be investigated further.

V. CORRELATION TECHNIQUES

In order to find the best fit of a control point chip in a CP neighborhood, a correlation surface is formed. This surface is generated after the enhancements are applied, and from its maximum (or minimum) we determine the best position of the chip. Several similarity measures were tried. The one used most was the normalized cross correlation(NCC), because of its sharp surface characteristics and its statistical nature. The NCC used is

$$\rho_{hk} = \frac{(N)^2 \sum_{i=1}^{N}\sum_{j=1}^{N} x_{ij}y_{h+i,k+j} - \bar{\mu}_x \cdot \bar{\mu}_y}{\sigma_x \cdot \sigma_y} \quad (5)$$

where

$$\bar{\mu}_x = \sum_{i=1}^{N}\sum_{j=1}^{N} x_{ij} \;,\; \bar{\mu}_y = \sum_{i=1}^{N}\sum_{j=1}^{N} y_{h+i,kj} \quad (6)$$

$$\sigma_x = \left[(N)^2 \sum_{i=1}^{N}\sum_{j=1}^{N} x_{ij}^2 - \left(\sum_{i=1}^{N}\sum_{j=1}^{N} x_{ij}\right)^2\right]^{\frac{1}{2}} \quad (7)$$

$$\sigma_y = \left[(N)^2 \sum_{i=1}^{N}\sum_{j=1}^{N} y_{h+i,k+j}^2 - \left(\sum_{1}^{N}\sum_{1}^{N} y_{h+i,k+j}\right)^2\right]^{\frac{1}{2}} \quad (8)$$

with a degenerate version for binary images.

To a lesser extent, some work was done using a Sequential Similarity Detection Algorithm of the form.

$$a_{hk} = \sum\sum \left| x_{ij} - y_{h+i,k+j} \right| . \quad (9)$$

Also suggested was an exponential similarity measure

$$e_{hk} = \frac{1}{N^2}\left(\sum\sum 2^{-L\left|x_{ij} - y_{h+i,k+j}\right|} \right) \quad (10)$$

where L is an expansion constant. This measure was not explored much due to lack of time. It is believed to have some interesting prospects since it is normalized to 1, is fairly simple and fast and points to an exponential type function for subpixel location.

Table 2. Enhancement Techniques Registration Success Ratios

Enhancemnt	Band2(5) 23 sets	Band3(6) 47 sets	Band4(7) 38 sets	Total sets
1. Hst+Gra+1σTH	48%	10%	50%	32%
2. Hst+Lap+15%TH	21	6	24	16
3. Hst+Lap	83	9	50	39
4. Hst+Gra	57	30	61	46
5. Lap+Hist	83	28	63	52
6. Gra+Hist	57	49	66	56
7. Hst Equal.	48	21	39	33
8. Lap+1σTH	39	10	34	25
9. Gra+15%TH	57	43	61	52
10. Laplacian	74	26	68	51
11. Gradient	52	51	63	56
12. NoEnhancemnt	52	26	45	28

VI. PIXEL REGISTRATION

Having created a correlation surface, we next locate its peak and determine whether a good registration was achived. To this end, a quick sort

routine is used to locate the 9 highest points of the surface. The 9 highest points are checked for clustering and the presence of a local maximum, which would be the second highest peak. Our criteria for a good registration was for the peak to be at least 3σ* units from the mean* of the surface heights and for the local maximum (if it exists) to be at least 2σ* units below the peak.

VII. SUBPIXEL REGISTRATION

Once the peak of a correlation surface has been accepted as a good registration, we need to find its location to subpixel accuracy. Several techniques to accomplish this were explored. Polynomials were fitted, to a neighboring array of the peak, by the method of least squares, and a center of gravity model was developed.

It should be noted that the subpixel location of the peak of a discrete correlation surface is totally dependent on the model used to describe the local surface. Furthermore, to our knowledge, there is no way of determining which of the many possible models are correct. For this reason, the goodness of a subpixel model was based on its stability under shifting and resampling of the surface.

Since it is believed that subpixel location on these statistical surfaces is a very local problem, our modeling was done on a 3x3 pixel neighborhood containing the peak as its center. In this neighborhood we know that the shape of the array (for a good registration) must be that of a cap. To model this cap with a polynomial we only need to look at a biquadratic of the form.

$$Z(x_1,x_2)=a_1x_1^2+a_2x_2^2+a_3x_1+a_4x_2+a_5. \qquad (11)$$

(Note the x_1x_2 term was omitted, since no saddle points are expected.)

The center of gravity model is given by

$$\bar{x}_i = \sum_{j=1}^{9} (w_j h_j x_{ij}) / \sum_{k=1}^{9} (w_k h_k) \qquad (12)$$

where W are weights and h are surface height differences from the smallest of the nine. Our results show, in Table 3, that the center of gravity model is more stable than the polynomials under our tests. We believe that the polynomial's instability arises from having errors in least squares fit of the same magnitude as the coefficients of the highest terms. If this is true, it would indicate that high order polynomials, which would have even smaller coefficients, would be less stable.

*Note: Robust statistics were used in determining the mean and σ.

Table 3.

Standard Deviations of Subpixel Location

Image Pair	L.I	C.C	L.I	C.C
7	0.20	0.24	0.19	0.52
8	0.05	0.06	0.27	0.28
9	0.13	0.10	0.36	0.39
10	0.05	0.06	0.43	0.40
11	0.10	0.12	0.33	0.38
12	0.06	0.10	0.34	0.39
Mean Deviation	0.098	0.11	0.32	0.39

VIII. CONCLUSIONS

The gradient enhancement is not the best in every situation but is also not weak in any and it gives the best overall results. Adding histogram equalization to the gradient improves it in some cases but has the drawback of shifting the data and costing computation time. The acceptance criteria of a registration should be explored and tuned in each individual system. The area of subpixel registration was a surprise and should be further investigated with other media and registration techniques. In particular, subpixel models to be used with a particular similarity measure should be studied.

ACKNOWLEDGMENTS

We wish to thank Mr. Lee Peters for his Programming Assistance and Dr. Ravindra Kumar for his help in compiling and analyzing the data. We also wish to thank Dr. Joan Brooks, Dr. Igore Levine, Dr. A. Singh and Mr. Stephen Lange for all their helpful comments.