Reprinted from

Symposium on

Machine Processing of

Remotely Sensed Data

June 21 - 23, 1977

The Laboratory for Applications of
Remote Sensing

Purdue University
West Lafayette
Indiana

IEEE Catalog No.
77CH1218-7 MPRSD

RECTIFICATION AND REGISTRATION OF DIGITAL IMAGES AND THE EFFECT OF CLOUD DETECTION

M. L. NACK
Computer Sciences Corporation

I. ABSTRACT

In previous image registration research, edge image correlation techniques were applied to Landsat-1 multispectral digital image data. Translations of edge images of an earlier area of interest were performed over a search area edge image at a later time to determine the location of the area of interest in the search area. Plots of the correlation surface indicated sharp peaks for successful searches and an accuracy of ±1 pixel or line. The following three new requirements were suggested: (1) because clouds or their shadows are are time-dependent image features, removal of cloud or cloud shadow information from the edge images before image correlation takes place; (2) accurate representation of geometric shapes of image features; (3) fractional pixel registration accuracy over subregions of the area of interest. Solutions to these problems were obtained. Distinguishing features of the geometric correction solution are the automatic location of the ground control points needed to determine correction parameters and the use of the geometric centers of edge image chips as the ground control points.

II. INTRODUCTION

This paper presents extensions of previous image registration research[1] which applied edge image correlation techniques to Landsat-1 multispectral digital image data. The success of this previous research resulted in the implementation of the Automatic Image Registration System (AIRS) by Computer Sciences Corporation (CSC) on a Sigma 3 computer for use by the Goddard Space Flight Center (GSFC) of the National Aeronautics and Space Administration (NASA). AIRS is used to automatically produce registered Landsat-1 and -2 imagery in a high-volume production mode and is achieving better than a 90-percent success rate with cloud-free imagery. This imagery is then sent to NASA's Lyndon B. Johnson Space Center (JSC) for multispectral classification in the Large Area Crop Inventory Experiment (LACIE).

In the current implementation of AIRS, nominal rotations and scale changes are performed on the image data using nearest neighbor resampling. Then, composite edge images are formed of an area of interest and a search area using bands 5 and 7 of the multispectral scanner (MSS). The edge image of the area of interest at one time is then translated over the search area edge image at a different time to determine whether the area of interest is totally contained in the search area and, if so, where. Plots of the correlation surface indicated sharp peaks for successful searches and an accuracy of ±1 pixel or line. This accuracy was corroborated in an independent study.[2] The advantages of the edge image registration technique were mentioned in an image registration comparison study.[3]

NASA suggested three problem areas for further investigation aimed at improving the quality of the AIRS product. First, clouds or their shadows are time-dependent image features, so that the removal of cloud or cloud shadow information from the edge images before image correlation takes place would then allow correlation of primarily time-independent features. Second, the geometric shapes of image features should be accurately represented to increase the probability of image correlations and also to produce a higher quality image registration product for user groups such as the JSC LACIE group. Finally, fractional pixel registration accuracy over subregions of the area of interest is desirable.

The solution to the cloud and cloud shadow problem involves simple thresholding techniques. Accurate representation of geometric shapes in Landsat images is achieved by performing sampling time delay corrections on all images due to sampling time delay between sensing image data from one line to the next in groups of six lines during a mirror sweep and the time delay between mirror sweeps of every six scan lines while the Earth rotates. Fractional pixel registration accuracy is achieved by performing subregion image correlations to generate general geometric correction parameters. Some of the details and results of these three solutions

are presented in Sections III, IV, and V, respectively, and a CSC[4] report contains further details.

III. CLOUDS AND THEIR SHADOWS

Clouds and their shadows are time-dependent image features. Information about these features should therefore be removed from images before attempting temporal correlation of two images of a geographic area taken at different times. The primary goal of the algorithm presented in this section is to remove edge pixels found in clouds or their shadows from the edge images and to produce a density of edge pixels D in the clear part of the image so that D is near D(MAX) but also $D \leq D(MAX)$. Clear pixels are those that are not covered by clouds or their shadows. Other goals of this algorithm were that the cloud and cloud shadow detection operation combine with the edge filter operation so that only a single pass of the image data through this combined operation is necessary. This goal aims at minimizing input/output (I/O) time on the small-core Sigma 3 AIRS computer. Another goal is that the algorithm be fast to minimize processing time on the AIRS production system.

The notation of Reference 1 will be used in this document whenever possible. Details of the data used in this study are presented in Tables 3 and 4 of Reference 1.

An edge image correlation should have an improved success rate by only correlating clear image pixels, and the probability of obtaining false correlation successes on the cloudy image features should be minimized. The current edge detection algorithm produces a high density of edges in the cloud features at the expense of producing edges on the more time-independent clear ground features.

An additional capability of the classified edge detection algorithm presented in this section is that while it classifies pixels into "edge" pixels and "not-edge" pixels, it further classifies not-edge pixels into cloud, cloud shadow, or clear pixels. The latter automatic classification can be used to automatically color cloud pixels black and cloud shadow pixels white to make them stand out in an image and to remove them from undergoing multispectral classification in the LACIE classification system.

The properties discussed in the preceding two paragraphs are illustrated in Figures 1 and 2. The technique for detecting cloud and cloud shadow pixels will be discussed later in this section and depends on cloud (T_C) and cloud shadow (T_S) thresholds which are shown as $T_S = 20$ and $T_C = 36$. A general comparison of the unclassified and classified images showed that though the cloud test performed quite well, the cloud shadow test performed poorly. The problem here was that as T_S was raised higher than 20 to classify correctly more cloud shadow pixels, this test then began to wrongly classify clear pixels.

The removal of the high density of edges in the clouds minimizes the chance of false edge image correlations resulting from two cloudy scenes. The removal of cloud and cloud shadow edges allows the threshold T of the combined or classified edge filter operation to decrease, and this results in (1) an obvious enhancement of the edges due to features in the clear part of the image, and (2) a corresponding improved probability of

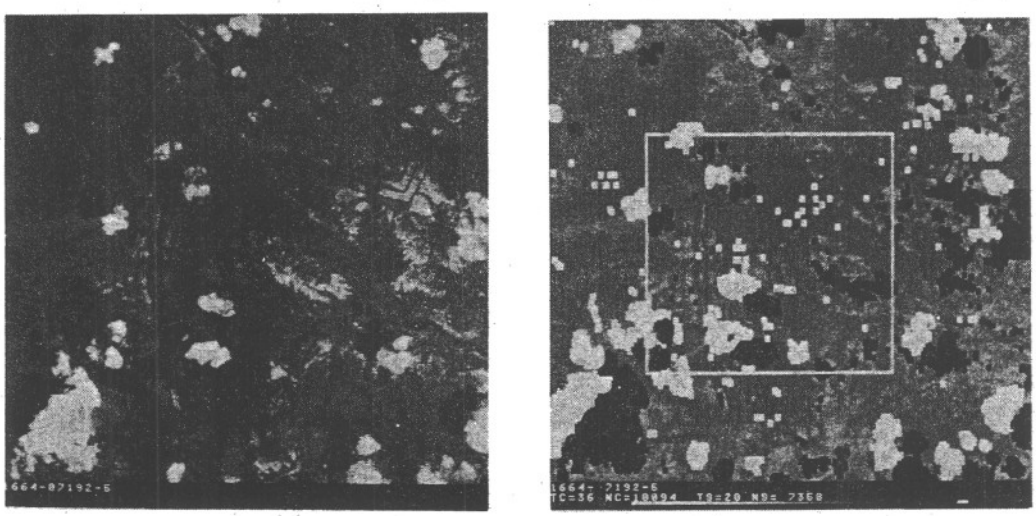


Figure 1. Subset of Landsat Scene 1664-7192-5 of Russia Containing Clouds, Cloud Shadows, and Snow.

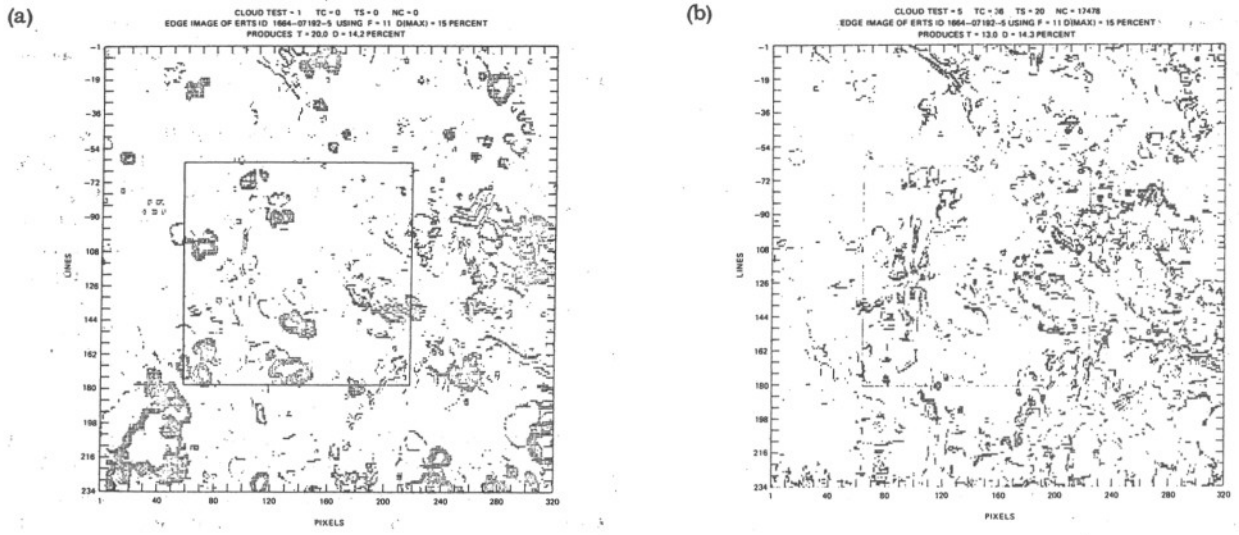


Figure 2. Edge Images of a Subset of Landsat Scene 1664-07192-5 (a) Without Cloud and Cloud Shadow Classification and (b) With Classification.

successfully correlating only the clear edge images of two scenes.

A subset of a Landsat scene of Russia is shown in Figure 1 which contains snow as well as clouds, plus some isolated black pixels. These black pixels which were classified as shadows appear to be some type of noise in the data. In Figure 1, the classification of clouds and their shadows appears quite accurate; however, many snow pixels are also classified as clouds. The edge images for this case in Figure 2 show the removal of edges due to clouds and their shadows and a lowering of the edge threshold from T = 20 to T = 13. The increase of edges in the clear part of this scene shows less apparent enhancement than in the other cases. The removal of edge pixels from Figure 2(a) due to the black noise pixels shown in the classified image of Figure 1 is also apparent in the edge image of Figure 2(b), e.g., the single black dot at (line 74, pixel 10) and cluster of 5 black dots about (line 84, pixel 30).

The algorithm for producing the classified edge-filtered image F_e^5 in band 5 for a general pixel e is shown in Figure 3. A successful test against cloud (T_c) or cloud shadow (T_s) thresholds assigns values of -1 and -2 for F_e^5 if pixel e is a cloud or a cloud shadow pixel, respectively. Only when both tests fail is the gradient of Reference 1 and Figure 3 computed and assigned to F_e^5. The absolute value of the difference of gray-level values I_i and I_j is denoted $D_{ij} = |I_i - I_j|$.

A histogram of the F^5 image would show the number of pixels in F^5 having the value L, or N_L^5, so that

$$N_c = \text{number of cloud pixels} = N_{-1}^5 \qquad (1)$$

$$N_s = \text{number of cloud shadow pixels} = N_{-2}^5 \qquad (2)$$

$$N_p = \text{number of pixels in } F^5 \qquad (3)$$

$$N_{pc} = N_p - N_c - N_s = \text{number of clear pixels in } F^5 \qquad (4)$$

$$n_L = 100\ (N_L/N_{pc}) \qquad (5)$$

where n_L is the percent of clear pixels in F^5 having the value L for L = 0, 1, 2, ...,. The values of n_L are used to determine the variable edge detection threshold, T, as described in Reference 1.

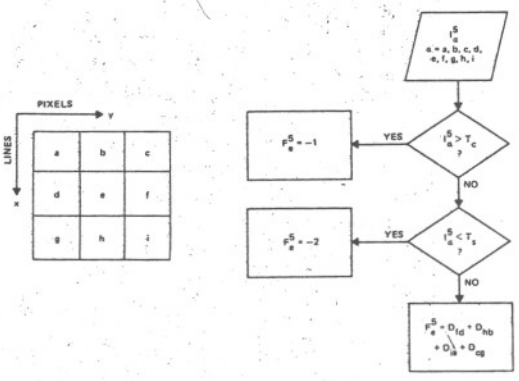


Figure 3. Flowchart for Computing Pixel Value in Classified Edge-Filtered Image F_e^5 From Pixel Values I_α^5 in Gray-Level Image for a General Pixel e.

This testing of all pixels within a 1-pixel radius of the test pixel e in Figure 3 is performed against T_c or T_s to allow the use of higher values of T_c or the lower values of T_s to correctly classify the interiors of clouds and their shadows and still correctly classify the borders of the cloud or cloud shadow. The values of T_c and T_s were determined by examining classified cloudy gray-level and edge images in comparison with unclassified images for a wide variety of cloud types and Sun angles as T_c and T_s are varied. Future work might determine a Sun angle dependence of T_c and T_s. Also, data-dependent thresholds, $T_c = \bar{I} + \tau_c\sigma$ and $T_s = \bar{I} - \tau_s\sigma$, might be preferable where $\bar{I}$ is the mean and σ the standard deviation of the signal, and τ_c and τ_s might be approximately equal to 2.

The classification is performed only in band 5 so that only one set of thresholds T_c and T_s need be determined, and a consistent classification is maintained in all bands. In band 7 if a cloud or cloud shadow pixel is detected ($F_e^5 < 0$), then that classification is also used in band 7 and the gradient is not computed, so that the processing time to compute a classified edge image decreases as the cloudiness of the data increases.

The results indicate that the algorithm decreases the number of edge pixels due to clouds and their shadows; increases the number of edge pixels in the clear portion of the image so that approximately D(MAX) of the clear image is edge pixels; classifies the output image into clear, cloud, and cloud shadow pixels; easily confuses snow and ice pixels with cloud pixels; and classifies data dropouts resulting in high or low signals or zero fill due to resampling as clouds or cloud shadows, and no edges are produced in the edge image due to these spurious discontinuities.

The determination of a reliable value of T_s proved quite difficult. A modification of the algorithm to allow the use of a sufficiently high value of T_s and yet not have pixels in a totally cloud-free image be classified as shadow pixels would be to require $N_c \geq 50$ before the cloud shadow test is turned on.

IV. ACCURATE REPRESENTATION OF GEOMETRIC SHAPES

This section analyzes the accurate representation of the geometric shapes of ground features at a given time and discusses resampling techniques. If geometric shapes of ground features of a given geographic area are represented accurately in images at two different times, temporal registration requires the mapping of the shapes at one time onto the corresponding shapes at the other time exposure. This mapping or image transformation is the topic of the next section. The geometric corrections or image transformations of both sections involve resampling the image in a new coordinate system. A general discussion of resampling and the nearest neighbor (NN) and cubic convolution (CC) resampling algorithms follow.

A. RESAMPLING TECHNIQUES

To obtain digital images, the continuous or analog image function I(x, y) is sampled on the discrete grid (x_i, y_j) to produce the discrete or digital image function $I_{ij} = I(x_i, y_j)$. It is often desirable to map image points in the (x, y) plane into identical image points in a new (u, v) plane. The reasons for performing this geometric transformation or mapping, denoted by g: (x, y) → (u, v) or (u, v) = g(x, y), might be to expand or contract an image or to rectify a wide variety of distortions present in I(x, y) to produce an undistorted image I(u, v). The desired transformed image is also usually a digital image $I_{kl} = I(u_k, v_l)$ sampled on the grid (u_k, v_l). Because the identical image point to (u_k, v_l) in the (x, y) plane is (x_k, y_l), determined by the inverse mapping g^{-1}: (u, v) → (x, y), where (x_k, y_l) in general does not fall on an (x_i, y_j) grid point, the general need to know I(x, y) for (x, y) not equal to an (x_i, y_j) grid point is generated; i.e., I(x, y) is resampled on the (x_k, y_l) grid generated by g^{-1}: $(u_k, v_l) \rightarrow (x_k, y_l)$.

The problem of selecting the optimum image interpolation function for a given image experiment is quite difficult, because consideration should be given to the interrelated parameters associated with the three main components of the experiment: the image target, the image receiving instrument, and the platform upon which the instrument is mounted. A general interpolation function could be written $I(x, y) = f[I(x_i, y_j), N(x, y)]$ where N(x, y) denotes the neighborhood of points (x_i, y_j) about (x, y) to be used in this particular function.

An unfortunate byproduct of resampling an image is geometric and/or radiometric image degradation. Though lower order techniques like NN geometrically degrade the image more than higher order techniques like CC, any interpolation of data involves some degradation. In addition, there is usually a tradeoff decision to be made between the increased processing time and core requirements and decreased degradation of higher order interpolation.

Because resampling results in degradation, repeated resampling operations on the same image should be avoided if possible.

The comparison of NN and CC resampling techniques has been thoroughly studied[5,6] and was recently documented by GSFC.[7,8] Because "nearest neighbor introduces a location error of up to 1/2 sample and line in each resampled pixel" (as reported in Section 3.7 of

Reference 7), the use of NN seems inconsistent with the accuracy goals discussed in Section V of this paper.

The goal of Section V is to register image shapes with fractional pixel accuracy. This will be achieved by fitting the discrete correlation surface P(l, p) with a continuous smooth surface P(x, y) in a 5-by-5 neighborhood of the peak of P(l, p) and then obtaining the peak of P(x, y) within a ±0.1-pixel (or line) error. Because an attempt to accurately represent geometric shapes should ideally be performed before temporal registration to allow maximum accuracy in the determination of discrete and continuous registration peaks, the jaggedness of using NN, which would be seen in the edge images, is also inconsistent with the goal of fractional pixel accuracy.

B. SEQUENTIAL GEOMETRIC CORRECTIONS

Consider a sequence of n geometric corrections to be performed on an image such that (x, y) is the original coordinate system upon which the image is sampled at the grid points (x_i, y_j) and the transformations from the corrected coordinate system (u, v) to (x, y) goes through n steps

$$(x, y) = g^1(u^1, v^1) \tag{6}$$

$$(u^1, v^1) = g^2(u^2, v^2) \tag{7}$$

$$\vdots$$

$$(u^{n-1}, v^{n-1}) = g^n(u^n, v^n) = g^n(u, v) \tag{8}$$

where $(u^n, v^n) = (u, v)$. If $g^1, g^2, \ldots, g^n$ are known, then it is preferable to form the combined transformation g

$$\begin{aligned}(x, y) &= g^1(u^1, v^1) \\ &= g^1[g^2(u^2, v^2)] \\ &\vdots \\ &= g^1\{g^2\{\ldots g^{n-1}[g^n(u^n, v^n)]\}\} \qquad (9) \\ &= g^1\{g^2\{\ldots g^{n-1}[g^n(u, v)]\}\} \\ &= (g^1 g^2 \ldots g^{n-1} g^n)(u, v) \\ &= g(u, v)\end{aligned}$$

$$g = g^1 g^2 \ldots g^{n-1} g^n \tag{10}$$

so that only one resampling operation is required on the (x, y) grid determined by the (u, v) grid, as opposed to n resampling operations on the (x, y), $(u^1, v^1), \ldots, (u^{n-1}, v^{n-1})$ grids. In general, geometric transformations do not commute, or $g^{12} \neq g^{21}$ where $g^{ij} = g^i g^j$.

In this paper, g^1 is described in Section IV. C and will represent the geometric transformation which results in an image $I_{g1} = O(g^1) \cdot I$, which has geometric shapes accurately represented. The notation of Reference 1 is used where the application of the geometric correction operation $O(g^1)$ on the uncorrected gray-level image I implies the transformation (x, y) = g^1(u, v) followed by resampling in the (x, y) plane on the (u, v) grid. Section V will describe g^2 or the geometric transformation which maps shapes of features in an image of an area of interest at time t_a onto the identical shapes in an image of the same area of interest at time t_b.

An example of the g^1 correction is shown in the images of Australia of Figure 4. The uncorrected image shows the vertical sawtooth distortion in all vertical shapes due to the sampling time delay between sensors and between groups of six lines during which the Earth rotates. The corrected image represents shapes accurately and uses CC resampling. The image has 234 lines and the removal of the 6-line sawtooth effect on the vertical diagonal linear features and on the curvy vertical road feature extending from the top to the bottom of the right part of the image is quite striking. The g^1 correction is the same correction which makes the standard Landsat film product have the parallelogram shape of the corrected image where the lower left portion of the parallellogram was truncated.

Because the g^1 correction produces images with accurate geometric shapes, a preferable way of achieving a final registered and rectified image from a search area SA image at time t_b, or $I(SA_b)$, and an area of interest AI image at time t_a, or $I(AI_a)$, is to perform the following processing steps: (1) form $I_{g1}(AI_a)$ and $I_{g1}(SA_b)$, where $I_{g1} = O(g^1) \cdot I$; (2) determine g^2 from automatically correlating ground control points (GCPs) or landmarks in $I_{g1}(AI_a)$ with the corresponding GCPs in $I_{g1}(SA_b)$; (3) form $g^{21} = g2g1$; and (4) produce $I_{g21}(AI_b) = O(g^{21}) \cdot I(SA_b)$.

The final registered pair of images, $I_{g1}(AI_a)$ and $I_{g21}(AI_b)$, are produced in steps 1 and 4 and require only one resampling operation to produce the corrected output images from the input images. A disadvantage of this procedure is that step 1 requires an additional resampling operation and corresponding processing time to form $I_{g1}(SA_b)$. The additional processing time to initially form $I_{g1}(SA_b)$ in step 1 was prohibitive for implementation on AIRS.

Figure 4. Subset of Landsat Image 1491-1003-5 of Australia (a) Uncorrected and (b) Corrected for Sampling Time Delays.

The following set of processing steps was decided to be investigated in this study; it is assumed that the image flaws that g^1 is trying to correct will average out and consequently will not affect the determination of g^2 from the unprocessed images: (1) determine g^2 from automatically correlating GCPs in $I(AI_a)$ with the corresponding GCPs in $I(SA_b)$; (2) form $g^{12} = g^1 g^2$; and (3) produce $I_{g12}(AI_b) = O(g^{12}) \cdot I(SA_b)$ and $I_{g1}(AI_a) = O(g^1) \cdot I(AI_a)$.

The advantages of this procedure are that (1) the initial resampling of $I(SA_b)$ is eliminated and (2) the final resampling is needed for only AI size images.

A disadvantage of both procedures is that they assume that the distortion between $I(SA_b)$ and $I(AI_a)$, which is to be measured in the determination of the g^2 mapping parameters, is small enough to not affect the ability initially to correlate $E(AI_a)$ with $E(SA_b)$. This should be checked by applying $O(g^2)$ to $I(SA_b)$ using known mapping parameters (e.g., rotations of 0.5 degree, 1.0 degree, and 1.5 degrees) and attempting a correlation which previously had succeeded with $I(AI_a)$. If this test is performed for autocorrelation when $t_a = t_b$, then this is a check also on the system's determination of g^2, since preprocessing $I(SA_b)$ by g^2 should cause the system to determine a $(g^2)^{-1}$ or the inverse correction (e.g., rotate back by -0.5 degree, -1.0 degree, and -1.5 degrees).

C. SAMPLING TIME DELAY CORRECTIONS

These corrections were initially addressed by Connel[9] and are discussed in detail in Section 6.8 of Reference 7 and shown in Figure 34 of that section. The sensor sampling time delay correction g^s depends on the sensor number $n = 1, 2, \ldots, 6$, and the Earth rotation sampling time delay correction g^e depends on the index $M = 0, 1, 2, \ldots$, which orders in a sequence groups of six scan lines. The g^1 correction depends on g^s and g^e, or $g^1 = g^e g^s$. In Figure 4, a g^1 corrected image of Australia is compared with the uncorrected image of the same area.

To determine n, M, and a given g^1 correction transformation, the image coordinate systems must be defined. The image coordinate systems (X, Y) for a full Landsat Scene S, (x, y) for SA extracted from S at (L_0, P_0), and (u, v) for AI extracted from SA at (l_0, p_0) are shown in Figure 5. To correct S, the scan line number X is expressed in terms of M and n

$$X = 6M + n \tag{11}$$

$$M = 0, 1, 2, \ldots \tag{12}$$

$$n = 1, 2, \ldots, 6 \tag{13}$$

so that

$$M = M(X) = \frac{1}{6}(X - 1) - \frac{1}{6}(n - 1) = \mathrm{INT}\left[\frac{1}{6}(X - 1)\right] \tag{14}$$

$$n = X - 6M \tag{15}$$

where INT(y) = largest integer $\leq y$. To correct SA, the equations

$$m = M(X) - M(L_0) \tag{16}$$

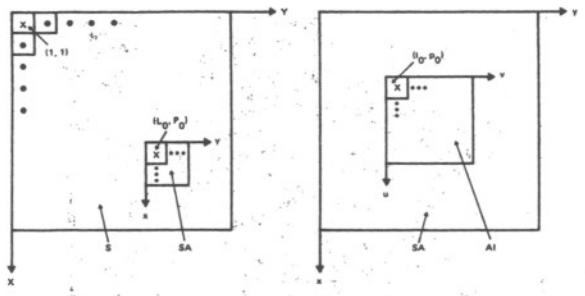


Figure 5. Image Coordinate Systems (X, Y) for a Full Landsat Scene S, (x, y) for a Search Area SA Extracted From S at (L_0, P_0), and (u, v) for an Area of Interest AI Extracted From SA at (l_0, p_0).

$$x = X - L_0 + 1 \tag{17}$$

are required.

The (u, v) coordinate system of AI will be assumed to describe the undistorted or corrected image of AI. The relationship between (x, y) and (u, v) if only the g^1 correction to SA is needed is

$$x = u + l_0 - 1 \tag{18}$$

$$y = v + p_0 - 1 - \left(\frac{22.0}{56.5}\right)\left(\frac{n-1}{5}\right) + \left(\frac{33.8}{56.5}\right)\cos(\text{latitude})\, m \tag{19}$$

where $g^1 = g^s$ if the m-dependent term is missing from Equation (19), $g^1 = g^e$ if the n-dependent term is missing from Equation (19), and $g^1 = g^{es}$ as given by Equations (18) and (19).

The application of the g^1 correction using CC resampling to a variety of images (1) produced a markedly improved accuracy in the representation of geometric shapes, (2) allows pairs of images to be more precisely registered, (3) removes the sawtooth blur to vertical shapes of features, and (4) allows more accurate multispectral and multitemporal image classification due to improved representation and registration of the borders of features to be classified.

V. FRACTIONAL PIXEL REGISTRATION ACCURACY

This section first discusses the process of image rectification which involves transforming a distorted image into an undistorted image. The ability to determine the transformation parameters with fractional pixel accuracy using edge image correlation techniques is then presented. This section closes with the presentation of images of Brazil and India which show areas of interest registered with fraction pixel accuracy.

A. INTRODUCTION TO IMAGE RECTIFICATION

The goal of image rectification is, given a distorted image I(x, y) and an undistorted image I(u, v) of the same area of interest, to determine the unique rubber sheet transformation which maps all (x, y) points recognizably identical with (u, v) points into these (u, v) points. In this application, I(u, v) is an image of a chosen AI, and I(x, y) is an image of an SA to be searched for containing AI, as shown in Figure 6. In practice, there is only a finite set of recognizably identical points or GCPs (u_α, v_α) and (x_α, y_α) for $\alpha = 1, \ldots, N$, where (u_α, v_α) and (x_α, y_α) are, in general, real numbers. Two distinguishing features of the algorithm presented in this section are (1) the automatic location of the GCPs and (2) the fact that the GCP is the geometric center of an edge image chip, as opposed to a particular recognizable point in the image.

The rubber sheet transformation can be visualized by sticking pins in I(x, y) at (x_α, y_α) and stretching, compressing, or twisting I(x, y) over I(u, v) so that the pins at (x_α, y_α) stick through I(x, y) over I(u, v) and into I(u, v) at (u_α, v_α). This determines a mapping g^2 of the (u, v) plane onto the (x, y) plane, denoted by $(x, y) = g^2(u, v)$, or $x = x(u, v)$ and $y = y(u, v)$, which is not unique but does satisfy the following GCP constraints: $x_\alpha = x(u_\alpha, v_\alpha)$ and $y_\alpha = y(u_\alpha, v_\alpha)$, for $\alpha = 1, \ldots, N$. These constraints force any selection of g^2 to map (u_α, v_α) into (x_α, y_α). To finally determine the correctly sampled image an evaluation of

$$I(x_k, y_l) = I[x(u_k, v_l), y(u_k, v_l)] \tag{20}$$

is required using a resampling technique of Section IV, because (x_k, y_l) are, in general, not grid points in the (x, y) plane of Figure 6.

In this approach of image rectification, five GCPs are defined in AI_a to be the centers of five subregions SR_α, for $\alpha = 1, 2, \ldots, 5$, of AI. The image sizes and locations of SA_b, AI_a, and SR_α for a sample location of AI_a found in SA_b are shown in Figure 6. The SR_α are located on the four corners of AI_a and at the center of AI_a. The coordinates of the upper left corners of SR_α in AI are denoted $(\bar{u}_\alpha, \bar{v}_\alpha)$. To determine (x_α, y_α) for $\alpha = 1, \ldots, 5$ in SA_b, SA_b must be searched for AI_a, as described in Reference 1. If AI_a is found in SA_b, this alignment of AI_a on SA_b is used to obtain likely first guesses $(x_{0\alpha}, y_{0\alpha})$ of the locations of the centers of SR_α in SA_b. Next, the edge images $E(SR_\alpha)$ are moved in a ±5-pixel (or line) search area about these guess locations (i.e., an 11-by-11 search area for finding SR_α in SA_b) and the five correlation probability arrays $P^\alpha(l, p)$ are calculated.

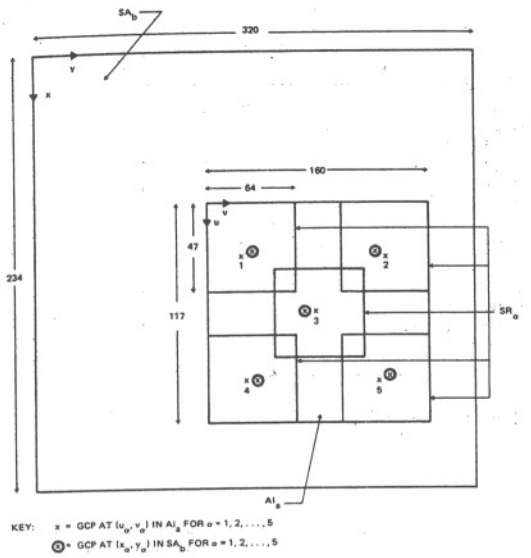


Figure 6. Image Sizes and GCP Locations.

The peak decision analysis of Reference 1 is then used to determine if unique peaks exist. If a peak does exist in $P^{\alpha}(l, p)$, then a 5-by-5 region about this peak will be fitted with a smooth surface $P^{\alpha}(x, y)$, and (x_{α}, y_{α}) will be determined as the peak of $P^{\alpha}(x, y)$ within ± 0.1 pixel (or line).

The peak decision analysis for $P^{\alpha}(l, p)$ is determined by calculating the mean, $\overline{P}_{\alpha}$, and standard deviation, σ_{α}, for the lowest 96 values of the $121 = 11 \times 11$ values of $P^{\alpha}(l, p)$. The existence test uses $\overline{P}_{\alpha} + 3\sigma_{\alpha}$, and the clustering test for uniqueness will then be performed on the top nine values of $P^{\alpha}(l, p)$. The selection of 3 as the multiple of σ_{α} and 9 as the number of top values to use for the peak existence and uniqueness tests were taken from Reference 1 based on results using larger images. These tests might have to be modified to be successful when using smaller subregion size images.

If a unique peak does not exist for SR_{α}, then the nominal position of (x_{α}, y_{α}) could be calculated using orbit and attitude data. An alternate approach which could increase the probability of finding a peak and which could use smaller subregions would be to decompose SR_{α} into four quarter size subregions $SR_{\alpha\beta}$ for $\beta = 1, 2, 3, 4$. Then edge image correlations could be performed on $SR_{\alpha\beta}$ with their subsets beneath them in SA_b, but in the order such that the number of edges in $SR_{\alpha\beta}$ plus its SA subset decreases. The greater number of edges in $SR_{\alpha\beta}$ and its subset in SA increases the chance of a successful correlation. Also, total processing time decreases in comparison to that of using the larger SR_{α} images if the first $SR_{\alpha\beta}$ correlation is a success.

B. AUTOMATIC GCP LOCATION

In this section, (x, y) will initially refer to the coordinates of the upper left corners where the SR_{α} from AI_a are found to be in SA_b. Later in this section, (x, y) will refer to GCP coordinates, and this is the case for the six pairs of images of Table 1 analyzed in this section. Table 1 shows for these six data sets the first guesses of the GCP positions $(x_{0\alpha}, y_{0\alpha})$, the GCP positions determined by the subregion edge image correlation technique described in this section (x_{α}, y_{α}), and the differences $\Delta x_{\alpha} = x_{0\alpha} - x_{\alpha}$ and $\Delta y_{\alpha} = y_{0\alpha} - y_{\alpha}$.

The results of calculating $P^{\alpha}(l, p)$ by correlating $E(SR_{\alpha})$ with $E(SA_b)$, the results of calculating where the upper left corners of SR_{α} should be if no rubber sheet distortion was present using

$$x_{0\alpha} = \overline{u}_{\alpha} + l_0 - 1 \tag{21}$$

$$y_{0\alpha} = \overline{v}_{\alpha} + p_0 - 1 \tag{22}$$

DATA SETS	LANDSAT ID OF		l_0	p_0	$\begin{Bmatrix} (x_{0\alpha}, y_{0\alpha}) \\ (x_{\alpha}, y_{\alpha}) \\ (\Delta x_{\alpha}, \Delta y_{\alpha}) \end{Bmatrix}$ WHERE $\begin{Bmatrix} \Delta x_{\alpha} = x_{0\alpha} - x_{\alpha} \\ \Delta y_{\alpha} = y_{0\alpha} - y_{\alpha} \end{Bmatrix}$ FOR $\alpha = 1, \ldots, 5$									
	AI (t_a)	SA (t_b)			1		2		3		4		5	
1. AUSTRALIA	1491–01334	1527–01331	43	68	67.0 66.6 0.4	100.5 100.2 0.3	67.0 66.0 1.0	196.5 198.2 −1.7	101.0 100.3 0.7	147.5 147.5 0.0	137.0 136.8 0.2	100.5 99.6 0.9	137.0 136.4 0.6	196.5 197.3 −0.8
2. BRAZIL	1105–12532	1537–12495	38	51	62.0 61.1 0.9	83.5 82.4 1.1	62.0 61.1 0.9	179.5 178.7 0.8	96.0 95.8 0.2	130.5 129.7 0.8	132.0 131.9 0.1	83.5 83.5 0.0	132.0 131.9 0.1	179.5 179.4 0.1
3. CHINA	1463–01574	1642–01483	65	63	89.0 88.1 0.9	95.5 96.7 −1.2	89.0 89.1 −0.1	191.5 192.1 −0.6	123.0 123.2 −0.2	142.5 142.4 0.1	159.0 158.6 0.4	95.5 94.3 1.2	159.0 159.0 0.0	191.5 191.0 0.5
4. CHINA	1642–01483	1661–01534	70	67	94.0 95.3 −1.3	99.5 98.7 0.8	94.0 94.0 0.0	195.5 194.6 0.9	128.0 128.3 −0.3	146.5 146.1 0.4	164.0 164.9 −0.9	99.5 100.1 −0.6	164.0 164.1 −0.1	195.5 196.1 −0.6
5. INDIA	1170–05023	1224–05030	18	124	42.0 41.7 0.3	156.5 156.4 0.1	42.0 41.2 0.8	252.5 251.5 1.0	76.0 75.3 0.7	203.5 203.3 0.2	112.0 111.9 0.1	156.5 156.5 0.0	112.0 111.5 0.5	252.5 251.6 0.9
6. U.S. (MONTANA)	1267–17404	1304–17461	67	66	91.0 91.4 −0.4	100.5 98.8 1.7	91.0 90.2 0.8	196.5 196.2 0.3	125.0 124.9 0.1	147.5 147.2 0.3	161.0 161.4 −0.4	100.5 101.1 −0.6	161.0 160.1 0.9	196.5 197.8 −1.3

Table 1. GCP Displacement Parameters for Six Data Sets.

FIG. NO.	DATA SET	α	$P^{\alpha}(l, p)$					$(x_{0\alpha}, y_{0\alpha})$	(x_{α}, y_{α})
7–a	2	4	l/p	49	50	51	52	53	(108, 51)	(107.9, 51.0)
			106	0.298	0.357	0.392	0.352	0.283		
			107	0.324	0.438	0.488	0.418	0.339		
			108	0.354	0.489	0.589	0.481	0.378		
			109	0.273	0.354	0.412	0.386	0.314		
			110	0.209	0.257	0.290	0.268	0.252		
7–b	5	5	l/p	217	218	219	220	221	(88, 220)	(87.5, 219.1)
			85	0.313	0.320	0.319	0.321	0.318		
			86	0.315	0.292	0.284	0.282	0.303		
			87	0.316	0.380	0.430	0.395	0.344		
			88	0.331	0.367	0.426	0.388	0.331		
			89	0.289	0.303	0.302	0.290	0.303		

Table 2. The Values of $P^{\alpha}(l, p)$, $x_{0\alpha}$, $y_{0\alpha}$) and (x_{α}, y_{α}).

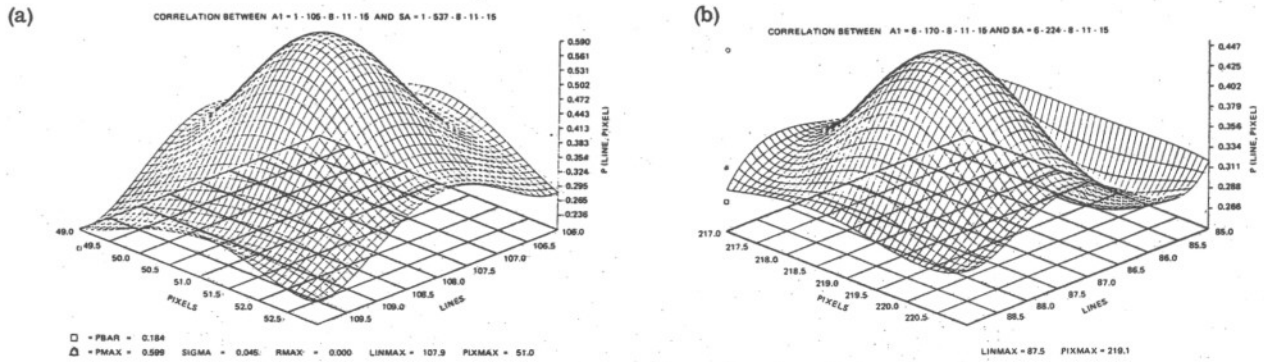


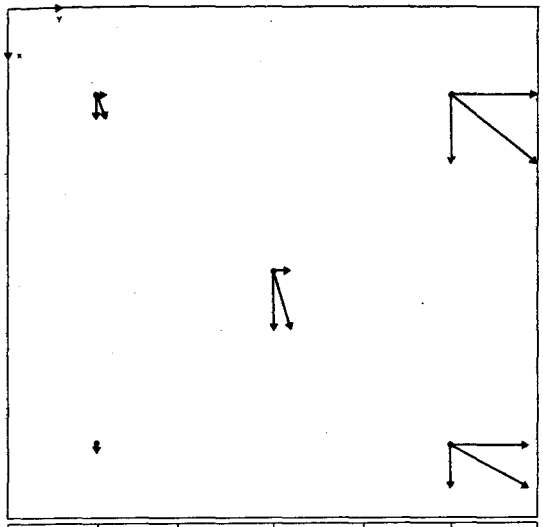
Figure 7. Correlation Surface Fit for (a) Data Set 2 and $\alpha = 4$ and (b) Data Set 5 and $\alpha = 5$.

and the location of the upper left corners obtained by fitting $P^{\alpha}(l, p)$ with a smooth surface is shown for two cases in Table 2 and Figure 7.

The technique used to fit the 25 points $P(x_i, y_j)$ centered about the peak of the discrete surface (l_1, p_1) is that of the Lagrange interpolating polynomial (e.g., see page 295 of Reference 7). This fitting procedure has the advantage of fitting all 25 points exactly using a fourth-degree polynomial in x and y, and the explicit nature of the fit allows easy implementation and fast computation. Three-dimensional plots of this P(x, y) surface were produced to allow examination of the fit, and they are shown in Figure 7. The ability of this surface to find values of (x_{α}, y_{α}) almost midway between sampling grid points is shown in Figure 7(b). Though Figures 7(a) and 7(b) both show well-defined correlation peaks, visual inspection of the Brazil data used for Figure 7(a) allowed manual correlation, whereas visual inspection of the India data used for Figure 7(b) defied manual correlation attempts due to the small irregular features in the India data. (See Reference 1 for the edge images of these two data sets and Figures 9 and 10 of the next section.)

This correlation surface fitting technique was applied to the comparison of the GCP positions (x_{α}, y_{α}) resulting from the application of this technique with the GCP positions, $x_{0\alpha} = u_{\alpha} + l_0 - 1$ and $y_{0\alpha} = v_{\alpha} + p_0 - 1$, resulting when no rubber sheet distortions are present.

The results are presented in Table 1 for six data sets, and the GCP displacements vectors for $\alpha = 1, 2, \ldots, 5$ are graphed in Figure 8. In the design of this technique to determine GCP positions using edge image chip correlation results, it was a source of concern that, due to the movement of image features like river banks in time, the GCP displacement vectors would display uncorrelated displacements of GCPs instead of displaying correlated GCP displacements due to variations in satellite orbit and attitude parameters in time as the satellite obtains images of the same AI at different times. The graphs of the results of all six data sets showed correlated displacement vectors. Figure 8 shows a clockwise rotation and shear displacement. The fact that the GCP displacement vectors show correlated distortion effects between temporal pairs of images of AI indicates that the technique of locating GCPs is performing satisfactorily. A quantitative interpretation of



α	1	2	3	4	5
q_α	18.32	0.9758	-5.566×10^{-2}	1.489×10^{-5}	2.256×10^{-4}
r_α	123.6	1.429×10^{-3}	1.011	0.0	-1.268×10^{-4}

Figure 8. GCP Displacement Vectors for Data Set 5 and the Resulting Mapping Coefficients.

the displacement vectors could be made using the results of the next subsection, which measure the magnitudes of rotation, scale change, and other distortions by determining mapping coefficients between the two images.

C. IMAGE MAPPING OPERATION

Given the location of the five GCPs (x_α, y_α) in what is assumed to be the distorted image of AI_a found in SA_b, an explicit mapping function g^2: $(u, v) \rightarrow (x, y)$ is needed to complete the image rectification and produce fractional pixel registration accuracy between the two images. This step assumes that a mapping function which maps five GCPs widely distributed in AI_a onto SA_b will also map with fractional pixel accuracy the remaining pixels of AI_a onto their corresponding pixels in SA_b.

The version of g^2 tested in this paper uses quadratic polynomials, as follows:

$$x = x(u, v) = q_1 + q_2 u + q_3 v + q_4 uv + q_5 u^2 \qquad (23)$$

$$y = y(u, v) = r_1 + r_2 u + r_3 v + r_4 uv + r_5 v^2 \qquad (24)$$

The five GCP constraints on g^2 can be written

$$x_\alpha = x_\alpha(u_\alpha, v_\alpha) = q_1 + u_\alpha q_2 + v_\alpha q_3 + (u_\alpha v_\alpha)\, q_4 + (u_\alpha^2)\, q_5 \qquad (25)$$

$$y_\alpha = y_\alpha(u_\alpha, v_\alpha) = r_1 + u_\alpha r_2 + v_\alpha r_3 + (u_\alpha v_\alpha)\, r_4 + (v_\alpha^2)\, r_5 \qquad (26)$$

for $\alpha = 1, 2, \ldots, 5$, or as matrix equations, $x = Qq$ and $y = Rr$, using the GCP vectors $x = x_\alpha$ and $y = y_\alpha$.

The solutions for the mapping coefficients of g^2 are obtained by inverting the Q and R matrices to give $q = Q^{-1}x$ and $r = R^{-1}y$, and this inversion need be performed only once because Q and R depend on (u_α, v_α), which are fixed coordinates in AI. The computation of Q^{-1} and R^{-1} could be performed once offline so that a small core production system would not require the core or processing time to compute Q^{-1} and R^{-1} for each set of data. To produce the rectified and resampled image of AI_a from SA_b now only requires resampling on an image the size of AI.

Because g^2 contains the operations of rotations and/or scale changes, the g^2 program module will have as an option the input of rotation and/or scale change parameters (q_α, r_α) for $\alpha = 1, 2, 3$ to preprocess I(SA). The need for this test option was discussed at the end of Section IV. B.

After determining g^2 from the SR_α correlations, a combined operation must be formed with g^1. The combined operation g^{12} is obtained by first using Equations (23) and (24) to perform $(x, y') = g^2(u, v)$ and then using the n- and m-dependent terms of Equation (19) to perform $(x, y) = g^1(x, y')$, because they only act on the y axis. The combined operation

$$(x, y) = g^{12}(u, v) = g^1[g^2(u, v)] = g^1(x, y') \qquad (27)$$

is given by

$$x = (u, v) = q_1 + q_2 u + q_3 v + q_4 uv + q_5 u^2 \qquad (28)$$

$$y = y(u, v) = r_1 + r_2 u + r_3 v + r_4 uv + r_5 v^2 - \left(\frac{22.0}{56.5}\right)\left(\frac{n-1}{5}\right) + \left(\frac{33.8}{56.5}\right) \cos(\text{latitude})\, m \qquad (29)$$

where the integers m and n for the Earth rotation and sensor delay corrections of g^1 are defined by Equations (11) through (16).

The mapping coefficients of q_α and r_α determined by the technique of this section for one of the six data

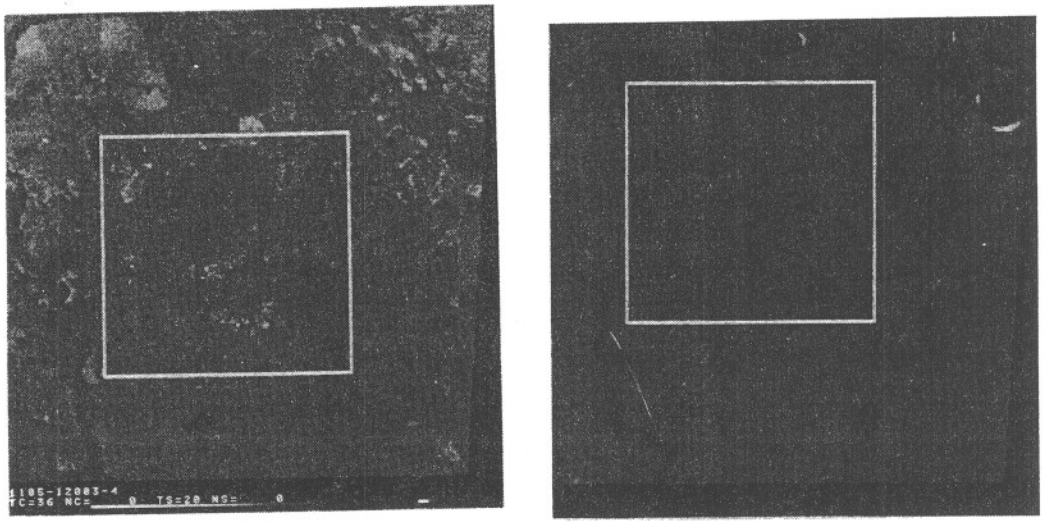


Figure 9. An Area of Interest From Brazil Registered With Fractional Pixel Accuracy.

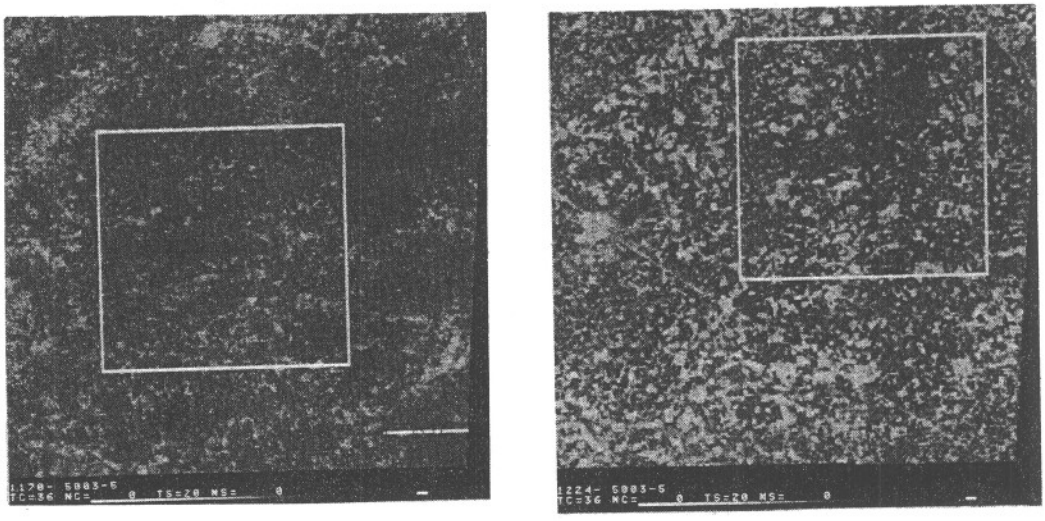


Figure 10. An Area of Interest From India Registered With Fractional Pixel Accuracy.

sets of Table 1 are shown in Figure 8. In Figures 9 and 10, these coefficients are used for data set 2 from Brazil and data set 5 from India to produce AIs registered with fractional pixel accuracy after the mapping of Equations (28) and (29) was performed and followed by CC resampling.

Figure 9 of Brazil uses band 4 to show AI_a and AI_b for a time separation of $\Delta t = t_a - t_b = 432$ days and displays high registration accuracy even though strong temporal changes took place in the images. Registered images of India are shown in Figure 10 using band 5 for images with a time separation of $\Delta t = 54$ days. The importance of the India data set is that attempts by skilled photograph interpreters to locate AI_a in SA_b failed when they could only use the image data in AI_a. However, after seeing the computer's result in locating AI_b using the edge image correlation technique of Reference 1, the photo interpreters (as well as the average viewer) were able to find enough recognizable features in both images to be convinced that accurate registration had been performed.

VI. CONCLUSIONS

The feasibility of efficiently classifying clouds and their shadows and then removing these pixels from the gray-level and edge images was demonstrated in Section III. The technique had the advantage of only requiring the determination of two thresholds in band 5 and then only performing the classification in band 5 so

that these results could then be used in the other bands. The ability to register data with clouds is greatly improved due to the enhancement of the clear part of the edge image, and the chance of false registration on clouds is minimized. Future improvements of the technique would require determining the thresholds from the data and only turning on the cloud shadow test after the scene was determined to be cloudy.

The superiority of using CC versus NN resampling to represent geometric shapes accurately (especially when fractional pixel accuracy is desired) was discussed in Section IV. The improved appearance of geometric shapes in Landsat imagery after sampling time delay corrections are performed was also confirmed. This seems to justify performing this correction before image correlation techniques are used to determine GCP mapping parameters for achieving fractional pixel registration accuracy, noting that the final products could still be produced with just one resampling operation if the sequence of the two geometric corrections for sampling time delay and GCP mapping were properly combined.

The ability to determine GCP positions automatically was demonstrated in Section V. The fitting of the discrete correlation surface with a smooth interpolating surface allows this GCP position determination with a tolerance of ± 0.1 pixel (or line). The ability of the mapping function to map all points or geometric shapes recognizable in both images accurately onto each other was demonstrated visually. Future work should be aimed at developing quantitative measures of registration accuracy to within a fraction of a pixel. One approach could be to take the final registered image products and use them as input to the subregion edge image correlation program. Ideally, the output should be zero length GCP displacement vectors.

VII. ACKNOWLEDGMENT

I would like to thank A. Anastas, H. Dennis, and E. Troy for their programming assistance; R. Krajewski for his analysis and programming assistance with the cloud and cloud shadow problem; C. Carlton for data handling assistance; and G. Grebowsky, Dr. R. White, and L. McQuinn for several helpful discussions.

VIII. REFERENCES

1. M. L. Nack, "Temporal Registration of Multispectral Digital Satellite Images Using Their Edge Images," Paper No. AAS 75-104, AAS/AIAA Astrodynamics Specialist Conference, Nassau, Bahamas, July 28-30, 1975

2. T. Kaneko and J. Engvall, "Evaluation of Image Registration Accuracy in the Large Area Crop Inventory Experiment," Symposium on Machine Processing of Remotely Sensed Data, June 1976, Purdue University, Indiana, p. 2B-2

3. M. Svedlow, C. D. McGillem, and P. E. Anuta, "Experimental Examination of Similarity Measures and Preprocessing Methods Used for Image Registration," Symposium on Machine Processing of Remotely Sensed Data, June 1976, Purdue University, Indiana, p. 4A-9

4. Computer Sciences Corporation, CSC/TM-76/6250, Final Report on Image Registration Research, M. L. Nack, November 1976

5. TRW Systems, 20634-6003-TU-00, Evaluation of Digital Correction Techniques for ERTS Images - Final Report, S. S. Rifman and D. M. McKinnon, March 1974

6. IBM Corporation, Final Report to GSFC Under Contract No. NAS5-21716, All Digital Precision Processing of ERTS Images, Ralph Bernstein, April 1975

7. Goddard Space Flight Center, Landsat Digital Image Rectification System Preliminary Documentation, Information Extraction Division, P. Van Wie, M. Stein, E. Pucinelli, and B. Fields, November 1975

8. --, X-931-76-101, A Landsat Digital Image Rectification System, P. Van Wie and M. Stein, May 1976

9. Ted Connel, "MSS Scan Line Anomaly," Memorandum to John Sos, February 14, 1973

10. E. Issacson and H. Keller, Analysis of Numerical Methods. New York: John Wiley and Sons, Inc.

Dr. Nack received his PhD in Theoretical Physics from Yeshiva University in January 1971. After 1-1/2 years of teaching physics in high school and junior high school in New York, he went on to a postdoctoral fellowship in theoretical elementary particle physics and atmospheric sciences at the University of Florida. In June 1974, he joined CSC to work in the areas of satellite image processing and atmospheric sciences under a NASA contract. At CSC he manages the Radiative Transfer and Climatology Section of the Remote Sensing Department.