

Spatial Registration of Multispectral and Multitemporal Digital Imagery Using Fast Fourier Transform Techniques

Paul E. Anuta, MEMBER, IEEE

Copyright © 1970 IEEE. Reprinted, with permission, from *IEEE Transactions on Geoscience Electronics*, Vol. GE-8, No. 4, October 1970.

This material is posted here with permission of the IEEE. Such permission of the IEEE does not in any way imply IEEE endorsement of any of Purdue University's products or services. Internal or personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution must be obtained from the IEEE by writing to pubs-permissions@ieee.org.

By choosing to view this document, you agree to all provisions of the copyright laws protecting it.

Spatial Registration of Multispectral and Multitemporal Digital Imagery Using Fast Fourier Transform Techniques

PAUL E. ANUTA, MEMBER, IEEE

Abstract—A system for spatial registration of digitized multispectral and multitemporal imagery is described. Multispectral imagery can be obtained from sources such as multilens cameras, multichannel optical-mechanical line scanners, or multiple vidicon systems which employ filters or other spectral separation techniques to sense selected portions of the spectrum. Spatial registration is required so that multidimensional analysis can be performed on contextually similar image elements from different wavelength bands and at different times. The general registration problem is discussed first; then the fast Fourier transform (FFT) technique for cross correlation of misregistered imagery to determine spatial distances is discussed in detail. A method of achieving translational, rotational, and scaling corrections between images is described. Results of correlation analysis of multispectral scanner imagery and digitized satellite photography is presented. Use of the system for registration of multispectral airborne line-scanner imagery and space photography is described. Application of the techniques to preprocessing of earth resources satellite imagery from systems such as the earth-resources technology satellite (ERTS) scanner and vidicon system is discussed in conclusion.

I. INTRODUCTION

REMOTE-SENSING technology is rapidly expanding into new forms of measurement. The data analysis and interpretation task is being automated to cope with the huge quantities of data generated by these advances. Photo interpretation has advanced from human analysis of panchromatic photography through similar manual analysis of color and color infrared photography to automatic computer analysis of multiband imagery. Human interpretation of multispectral photography which contains the spatial dimension and two or more spectral bands is cumbersome due to the difficulty of simultaneously observing the characteristics of multiple representations of a scene. Color and color infrared film greatly increases the information content of a single picture by introducing the spectral dimension into a single pictorial representation; however, the human speed and judgement factor is still a limitation in this form of analysis. The magnitude of the analysis task and the limited range of spectral response of film have led to the development of automatic data collection and analysis systems which can sense

and process large quantities of data from many measurement dimensions. The key requirement of an automatic multidimensional analysis system is the availability of a set of congruent measurements for each resolution element in the image. Multiple measurements from each image resolution element offer a means of improving the accuracy of recognition of the properties of the surface of the scene over that attainable using one dimension. Measurements of reflectance and radiance from microwave, thermal, and reflective infrared, through the visible wavelengths and into the ultraviolet region, can be utilized for analysis of each image point if congruence of these measurements can be achieved.

Remote-sensing measurement and analysis techniques are commonly classified into spectral, spatial, and temporal methods. Polarization measurement is often considered to be a fourth method. The spectral class was discussed previously. Use of the temporal or time-varying dimension also requires that image congruence be achieved so that automatic analysis can be carried out using this dimension. The time-varying properties of spectral and spatial features in a scene are included in the temporal dimension, and image registration is required in either case. Misregistration results from the inability of the sensing system to produce congruenced data due to design characteristics or the fact that the sensors are separated in space and time such that spatial alignment of the sensor is impractical or impossible. Geometric distortion, scale differences, and look angle effects can all combine to produce misregistration.

A digital system has been developed at the Laboratory for Applications of Remote Sensing (LARS), Purdue University West LaFayette, Ind., for spatial registration of multispectral-multitemporal imagery so that research into the usefulness of various measurement dimensions could be performed. The sources of multi-imagery considered by LARS are airborne multispectral optical-mechanical line scanners and multiband photography. However, the results of the work described are applicable to imagery from any source which produces multiple images of the same scene in different wavelength bands or at different times and in different wavelength bands. LARS is concerned with the development and application of remote-sensing technology to all aspects of earth-resources survey. Our work is

Manuscript received May 1, 1970. This work was supported by the U. S. Department of Agriculture under Contracts 12-14-100-9549 (20) and 12-14-100-10292 (20). This paper was presented at the 1970 IEEE International Geoscience Electronics Symposium, Washington, D. C., April 14-17.

The author is with the Laboratory for Applications of Remote Sensing, Purdue University, West Lafayette, Ind. 47906.

concentrated on development of digital multispectral and multitemporal statistical pattern-recognition methods which can identify agricultural, geological, and other features accurately and rapidly [1]. The work of Steiner [2] is an example of the use of the time dimension for pattern recognition. The digital approach was chosen by LARS to enable maximum flexibility in algorithm-development and data-analysis activities. LARS is an interdisciplinary project which includes engineers and scientists from the fields of electrical engineering, computer science, agronomy, soil science, geology, civil engineering, forestry, natural resource conservation, and others. Creative interaction between persons in these diverse areas has proven to be most easily achieved through user-oriented computer programs and a high-speed general-purpose digital computer. This form of interface allows persons with little computer or engineering knowledge to utilize the algorithms and data produced by the engineering technologists with very little training of the users. Also, testing and evaluation of data-handling and analysis schemes is easily carried out through software development rather than iterative hardware design and implementation. The decision to study image registration using sequential digital techniques was based on the success experienced using similar techniques in other areas of data-handling and analysis research at LARS.

The general multi-image registration problem is discussed first; then correlation techniques and problems are discussed. The use of the fast Fourier transform (FFT) for high-speed correlation of image arrays is discussed in detail; then the results of correlation analysis of test imagery from agricultural areas are presented. The use of enhancement techniques to improve the correlation accuracy of relatively uncorrelated wavelength bands is discussed next, and a description of the current LARS imagery registration system is presented. Application of these techniques and results to registration of aircraft-scanner imagery and Apollo 9 multispectral photography is discussed in conclusion.

II. IMAGERY REGISTRATION DEFINITIONS

Multiple images of the same scene are in registration if contextually coincident resolution elements in different images are uniquely addressable by one coordinate pair and an image pointer. A digital image is a two-dimensional array of numbers which represents the characteristics of a real image at discrete points. Digital image registration considerations include the sampling rate, sample point geometry, quantization effects, and other factors, as well as the geometric alignment of multiple images. Fig. 1 depicts the registration requirement for a three-member digital image set. The shaded resolution element represents an image point from the same scene point, and the coordinates (i, j) locate that element in each image. Spatial misregistration of multiple images can take the form of translational, rotational, and scale differences between image pairs.

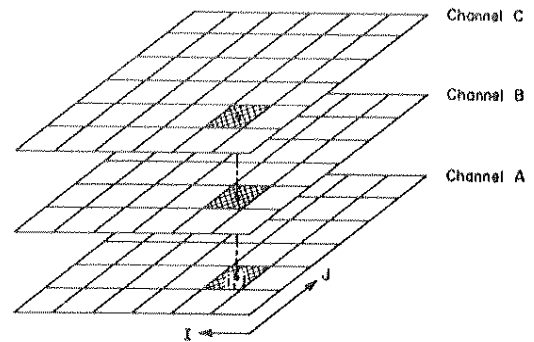


Fig. 1. Multi-image registration requirement. Resolution elements in registered multi-image are addressable by one coordinate pair and a channel pointer.

Multispectral and multitemporal images of the same scene can have markedly different characteristics. Although the context is the same in each of the images, the reflectance or emissivity of individual scene points may be totally different. Fig. 2 is an example of multispectral scanner imagery in 14 wavelength bands from $0.32 \mu\text{m}$ to $14 \mu\text{m}$ reproduced in image form on a computer line printer. The differences in spectral response can be seen by examining a particular scene area in all the bands. Multitemporal images may depict a scene which has undergone climatic or cultural changes such as seasonal changes, plowing, harvesting, or cloud-cover intrusion. A form of rotational misregistration in the multitemporal case can be seen in Fig. 3 which contains a computer printout of multispectral scanner imagery taken in June and August, 1969. The scanning aircraft flew with a large yaw angle in June, due to crosswinds, which caused the unstabilized scanner to produce the skewed imagery seen in the left photo. The August data from the same area is nearly geometrically correct. This data must be registered if it is to be used by a multitemporal pattern-recognition system. The general registration problem is thus one of determining the location of matching context points in multiple images and alteration of the geometric relationships of the images such that the registration of each context point is achieved.

The registration process is divided into three phases which aid in studying the problem; enhancement, correlation, and overlay. Enhancement refers to the preprocessing necessary to improve the accuracy of registration; correlation is the process of determining the location of matching context points; and overlay is the geometric transformation process which produces the registered imagery. A digital multi-image is defined as a set of $M \times N$ -point digital images or pictures P^k , where $k=1, \dots, K$,

$$P^k = \{x_{ij}^k\} \quad , i = 1, \dots, M; \quad j = 1, \dots, N.$$

One of the K multi-images q is selected as a reference or context image, and the other $K-1$ members of the set are registered with respect to P^q . The crux of the registration process is determining the spatial distance d_{ij}^k between each point in each of the $K-1$ images and the corresponding point in the context image. The distance

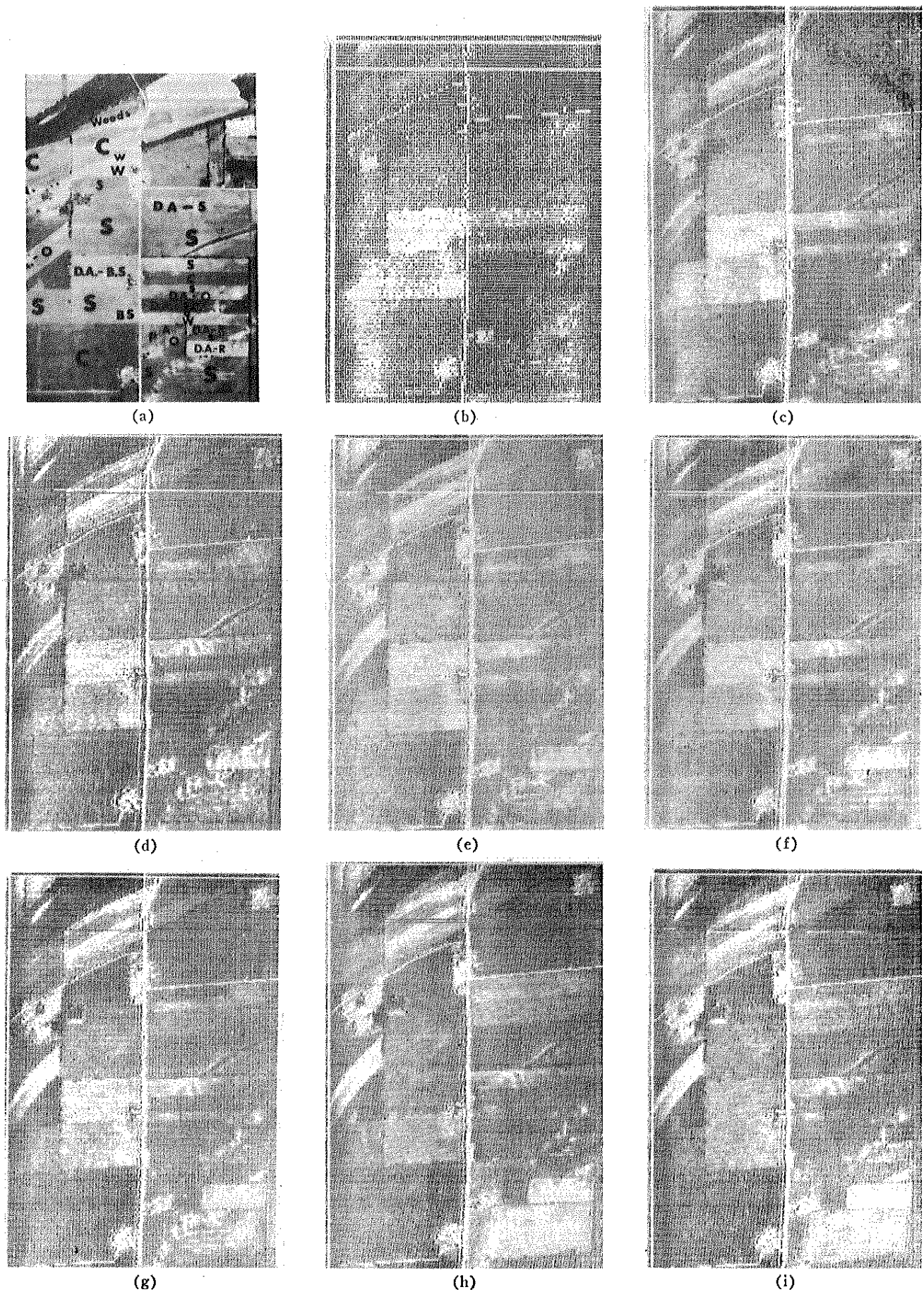


Fig. 2. Example of scanner imagery from 14 bands. (a) Aerial photo. (b) 0.32-0.38 μm . (c) 0.40-0.44 μm . (d) 0.44-0.46 μm . (e) 0.46-0.48 μm . (f) 0.48-0.50 μm . (g) 0.50-0.52 μm . (h) 0.52-0.55 μm . (i) 0.55-0.58 μm .

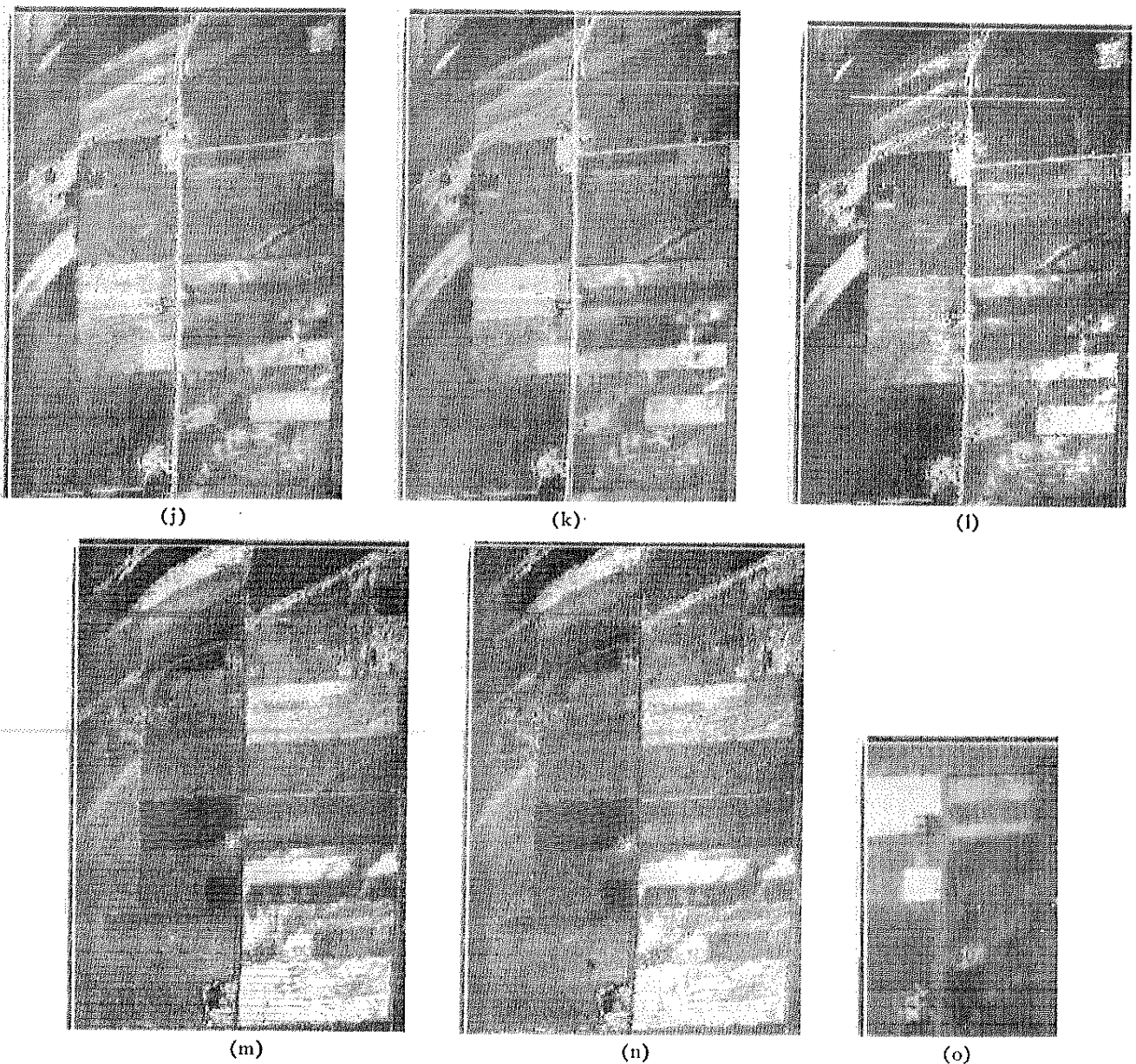


Fig. 2. Example of scanner imagery from 14 bands. (j) 0.58–0.62 μm . (k) 0.62–0.66 μm . (l) 0.66–0.72 μm . (m) 0.72–0.80 μm . (n) 0.80–1.0 μm . (o) 8–14 μm .

is defined in terms of translation, rotation, and scale differences. The general spatial distance for the image k -tuple on a point basis consists of three quantities.

1) Translational distance: $d_{T_{ij}^k} = (\Delta I_{ij}^k, \Delta J_{ij}^k)$ is an ordered pair of real numbers giving the translation on two orthogonal axes which point x_{ij}^q would undergo to be aligned with the corresponding point in image P^k .

2) Rotational distance: $d_{R_{ij}^k}$ is the angular rotation a neighborhood in image P^k would have to undergo to have the same angular orientation as the corresponding point in the context picture. Note that this definition is ambiguous for a single picture point and is assumed to refer to a small image neighborhood around the point x_{ij}^k .

3) Scale distance: $d_{s_{ij}^k}$ is the scale factor of the P^k picture with respect to P^q at point p_{ij}^k . It is an ordered pair giving the scale factor on two orthogonal axes. The scale factor is defined as the ratio of the incremental

image distance represented by a set of points in P^q to the distance represented by the same size set in P^k .

An additional problem encountered in digital imagery is the situation in which a picture point may not exist at the same context point in all images due to differences in sampling rate and sample reference points. This condition requires either resampling of the original image or interpolation between available image points to obtain a sample at the desired position. Also, differences in the size, shape, and values of the image elements in different bands constitute a registration error which must be corrected before the multi-image can be considered to be in accurate registration.

III. IMAGE CORRELATION

The core of the LARS registration system is an array correlator which computes the correlation coefficient of two image arrays for a set of juxtapositions in two

dimensions. The correlation is carried out by numerical operations on a sequential digital computer. This approach is in contrast to the more common method using electrooptical techniques in which analog voltages from an optical sensor are correlated. Since a digital picture is composed of discrete brightness points each having a specific numerical value, the correlation function is represented by a double summation over the points being correlated. The two dimensional discrete correlation function [3] for two pictures X and Y is

$$\phi(D) = \phi(k, l) = \frac{\sum_{i=1}^N \sum_{j=1}^N x(i+k+s, j+l+s) y(i, j)}{\sum_{i=1}^N \sum_{j=1}^N x^2(i+k, j+l) \sum_{i=1}^N \sum_{j=1}^N y^2(i, j)} \quad (1)$$

where

x_{ij}, y_{ij} points from pictures X and Y corrected to have zero mean. Picture X is of size M by M , Y is N by N , $N < M$.

k, l shift variables, $k, l = -s, \dots, s$, with $s = (M - N)/2$.

$D = (k, l)$ translational distance vector from the reference point.

This formula expresses the correlation on a scale from -1 to 1 where $\phi = 1$ indicates perfect positive correlation and $\phi = -1$ indicates perfect negative correlation. The expression assumes that the two pictures X and Y have the same scale, rotational alignment, and distortions. This is not the case in general, and these other degrees of freedom must be included in the correlation process. Rotational and scaling search processes must be carried out in addition to translational search if the angular and scale differences are severe. This costly search procedure can be simplified if initial checkpoints can be obtained before correlation. These checkpoints can be used to determine the preprocessing scale factor and rotation so that when correlation is carried out a reliable image match can be found. The scale factor and rotation are then updated from the translational correlation function results.

This approach was implemented at LARS early in the registration study. The search mode was used for each correlation operation, and for large search rectangles on the order of 10 to 20 points square the computation time was excessive. The track mode was unfeasible at the time since the correlation was not carried out on adjacent or nearby points and the variation of lock-on points over 100- to 200-image-point intervals tended to be large. Instead of improving the search-track method of correlation to improve processing speed, the use of the recently developed FFT for correlation was investigated.

The convolution theorem of Fourier analysis states that convolution in the time or space domain is equivalent to multiplication in the frequency domain. Since

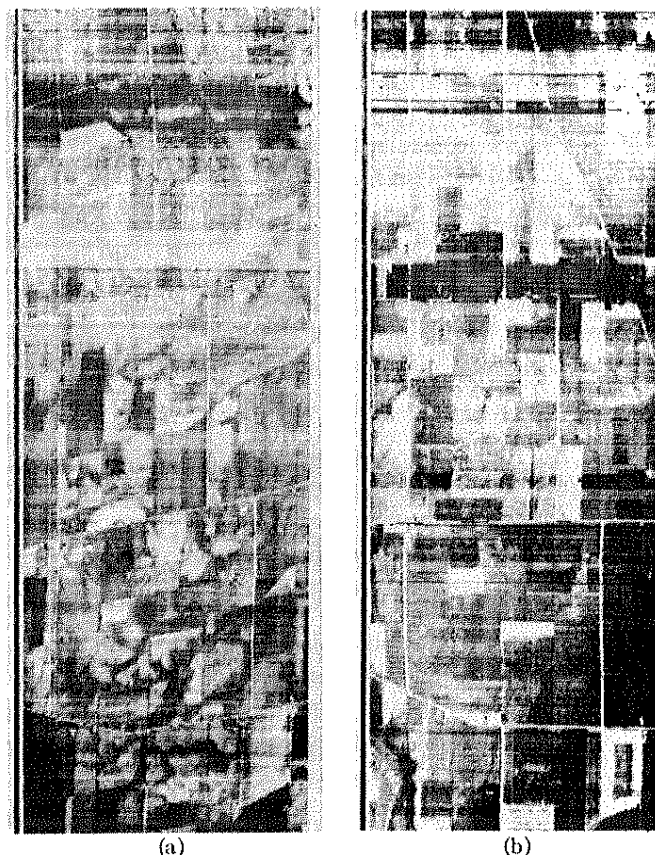


Fig. 3. Example of rotational distortion in multispectral airborne scanner imagery. (a) June, 1969; 11.7° yaw distortion. (b) August, 1969; 0° yaw distortion. Flightline: PF24; altitude: 5000 feet; band: $0.66-0.72 \mu\text{m}$.

correlation is a form of convolution, an alternate method of computing the correlation function thus exists. The classical numerical integration method of computing the Fourier transform is as time consuming as direct evaluation, and the method alone offers no great advantage. With the publication of an algorithm [4] for computing the Fourier transform which is orders of magnitude faster than the conventional method, a fast means of computing the correlation function became available. The time required to compute an n -point transform using the FFT varies as $n \log n$ instead of approximately n^2 for conventional numerical integration evaluation. Therefore, the transform method was implemented in the imagery registration system.

Certain problems unique to the use of the finite Fourier transform exist, and the solutions to them are outlined here. Convolution in one dimension is expressed mathematically as

$$C(k) = \int_{-\infty}^{\infty} x(t)y(k-t) dt = FT^{-1}[X(f)Y(f)] \quad (2)$$

where

$x, y(t)$ two time functions
 $C(k)$ the convolution between two time functions shifted by k units with respect to each other

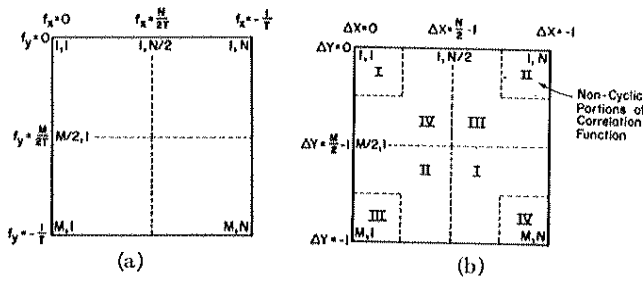


Fig. 4. Two-dimensional Fourier transform and correlation function structure. (a) Location of harmonic components of a 2-dimensional discrete Fourier transform. (b) Location of components of 2-dimensional correlation function computed by discrete Fourier transform.

$X, Y(f)$ the Fourier transforms of the two time functions; the transform is a function of the frequency variable f
 FT^{-1} signifies the inverse Fourier transform operation.

Cross correlation is defined in the same way as convolution except that one of the two functions is not reversed but is simply shifted on its axis. The correlation function is expressed as

$$\phi(k) = \int_{-\infty}^{\infty} x(t)y(t-k) dt = FT^{-1}[X(f)Y^*(f)] \quad (3)$$

where

- $\phi(k)$ the correlation function of the shift variable k
- $Y^*(f)$ the complex conjugate of the Fourier transform of $y(t)$.

Direct application of the preceding expression to computation of the correlation function of discrete data using the discrete Fourier transform leads to problems which may cause erroneous results. The Fourier transform algorithm computes the discrete N -term Fourier series of an N -point function. The location of the harmonic components and convolution values in the transform array is shown in Fig. 4. Inherent in the operation of the transform algorithm is the assumption that the function being transformed is periodic. The resulting N -point transform is also periodic. The result of the application of the preceding expression in the discrete case is a cyclical convolution function which is not the desired result. An example of the desired result compared to the cyclic result is shown in Fig. 5. This problem can be alleviated by increasing the size of the transformed array and including zero values for the range of shift desired. For $x(k)$ defined at M discrete points (M even) $k=0, 1, \dots, M-1$ and $y(k)$ defined at $N < M$ discrete points (N even) centered in an array of size M , let $y(k)=0$ for $k=0, 1, \dots, [(M-N)/2]-1$ and $k=(M+N)/2, \dots, M-1$. Then executing the correlation process using the transform will give the correct result for a shift of $(N-M)/2$ points in each direction.

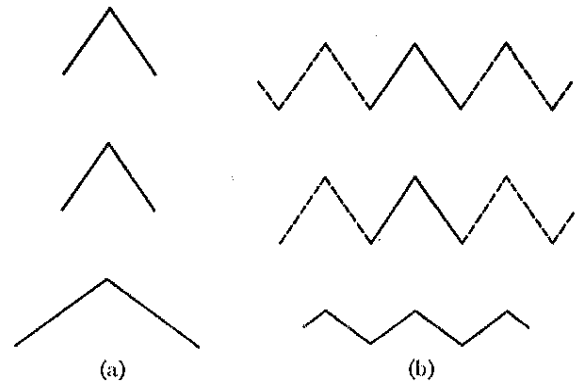


Fig. 5. Example of cyclical correlation problem. (a) Desired correlation function. (b) Cyclical correlation resulting from direct application of convolution theorem using the discrete Fourier transform.

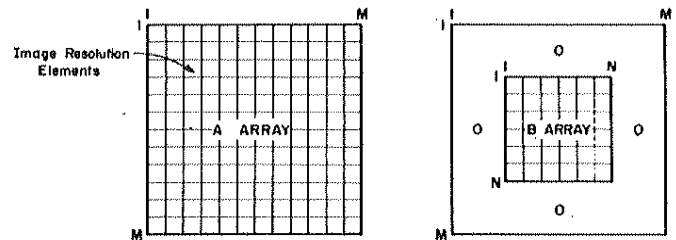


Fig. 6. Array structure for avoidance of cyclic correlation.

The correlation function

$$\phi(k) = \sum_{i=0}^{M-1} x(i)y(i+k), \quad k=0, \pm 1, \dots, \pm \frac{M-N}{2} \quad (4)$$

is thus computed using the FFT by multiplying the M -point transforms of the x array and the y array constructed with $M-N$ 0's added as follows

$$\phi(k) = FT^{-1}[X(f)Y^*(f)], \quad k=0, 1, \dots, M-1. \quad (5)$$

The $k=0$ point is the correlation for no shift, $k=1$ is for a 1-point shift in the positive direction and so on up to $(M-N)/2$ points of shift in one direction. The $(M+N)/2$ point is the correlation for maximum shift in the opposite direction and the $k=M-1$ value is a 1-point shift in the negative direction. This split is due to the cyclic property of the transform, and behavior of this type must be accounted for in a system which uses the transform technique. The values of $\phi(k)$ for $k = [(M-N)/2] + 1$ up to $k = [(M+N)/2] - 1$ are invalid and are not used. They represent the correlation of y shifted such that the values of $y(i)$ for $i > k$ are wrapped around the end of x and are being correlated with $x(0), x(1), \dots$, which is meaningless in most cases. This picture is changed if the 0's are included in the y array at different points. It can be stated in general that the only valid correlation function points are those with the same index values as the 0 points in the $y(i)$ function.

The cyclic convolution elimination problem for the two-dimensional case is solved in the same manner as for the one-dimensional case. The y points must be

surrounded by 0's on four sides to fill it out to the size of the larger x array, as shown in Fig. 6. The valid correlation points are identified in the same manner as for the one-dimensional case except that the four quadrants of the valid correlation function are at the four corners of the total cyclical correlation function. Specifically, let the x array be of size M by M [$x(i, j); i=0, \dots, M-1; j=0, 1, \dots, N-1$] and y' be N by N [$y'(i, j); i=0, \dots, N-1; j=0, \dots, M-1$] with $N < M$ (M and N even). The N by N points are assumed to be in the center of an M by M -size array, and this array is padded out with 0's such that

$$y(i, j) = 0, \quad i, j = 0, 1, \dots, \frac{M-N}{2} \quad \text{and} \quad i, j = \frac{M+N}{2}, \dots, M-1$$

$$y(i, j) = y'(i-s, j-s), \quad i, j = \frac{M-N}{2} + 1, \dots, \frac{M+N}{2} - 1; \quad s = \frac{M-N}{2} \quad (6)$$

where $y(i, j)$ is an M by M array.

This is the same basic format as for the one-dimensional case. Similarly, the valid correlation function points lie in the first $(M-N)/2$ points in each corner of the two-dimensional square result array.

In order to prevent data sets with large average values from "swamping" the correlation by effectively introducing a large square pulse into the data, the average value of the y data points is removed before padding in the 0's. The average value of x is also removed to minimize the magnitude of the correlation function. The root sum of the squares for the arrays is also computed and used to normalize the correlation function. A step-by-step account of the operations necessary for two dimensional correlation is now described.

1) Select the size in image points of the area to be covered by the correlation integral. Assuming it is

of squares is computed for x and divided into each point of the array. The same is done for y . The y set is placed in the center of the M by M array which is padded out with 0's as defined previously.

3) The M by M Fourier transform of x and y is computed using the FFT algorithm HARM (SHARE program SDA 3425 or similar versions). The complex conjugate of the Y transform is taken; the X and Y transforms are multiplied and the inverse transform is taken which produces the total M by M cyclical correlation function with the valid points at the corners of the array. Mathematically these operations are expressed

as follows. The two-dimensional transformation is

$$X(f, g) = \frac{1}{M^2} \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} x(i, j) W_M^{-f i} W_M^{-g j},$$

$$W_M = e^{2\pi\sqrt{-1}/M}. \quad (7)$$

$Y(f, g)$ is computed in the same manner.

Then

$$\phi(k, l) = \sum_{f=0}^{M-1} \sum_{g=0}^{M-1} [X(f, g) Y^*(f, g)] W_M^{k f} W_M^{l g} \quad (8)$$

where * indicates the complex conjugate.

4) The resulting total correlation function $\phi(k, l)$ is partitioned and the quadrants are interchanged as shown in the right of Fig. 6 to place the 0-shift point in the center of the $2\Delta+1$ by $2\Delta+1$ two-dimensional correlation function as follows

$$\psi(k + \Delta, l + \Delta) = \phi(k, l), \quad k, l = 0, 1, \dots, \Delta; \quad \left(\Delta = \frac{M-N}{2} \right)$$

$$\psi(k + \Delta, l - M + \Delta) = \phi(k, l), \quad k = 1, \dots, \Delta; \quad l = M - \Delta - 1, \dots, M - 1$$

$$\psi(k - M + \Delta + 1, l + \Delta) = \phi(k, l), \quad k = M - \Delta - 1, \dots, M - 1; \quad l = 0, \dots, \Delta$$

$$\psi(k - M + \Delta + 1, l - M + \Delta + 1) = \phi(k, l), \quad k, l = M - \Delta - 1, \dots, M - 1. \quad (9)$$

square, let this value be N (N even). (The rectangular case is a trivial extension of the square case.) Next select the maximum shift of one array with respect to the other. Let this be Δ . The correlation function will then have $2\Delta+1$ points in each direction, plus and minus Δ and 1 for 0-shift.

2) Step 1) defines the necessary size for the base or x array. It is $N+2\Delta$ square, and this is called M . The average of the M by M x array is removed, and the average of the N by N y array is removed. The root sum

The function ψ is a $2\Delta+1$ by $2\Delta+1$ array of correlation values with the 0-shift point in the center. This array is the output of the correlation routine. An example of a typical correlation function for highly correlated imagery is presented in Fig. 7.

A core storage saving scheme can be used when employing the FFT routine for transforming real data [5]. The transform algorithm is written to be able to process complex data; thus each input data point is a double computer word. To perform correlation, two separate

arrays must be transformed, i.e., the x and y data sets. The total number of data storage words required for the M by M transform is thus $2 \cdot 2 \cdot M^2$. The core saving method is based on the fact that the real part of the Fourier transform of real data is even about the 0 frequency point and the imaginary part is odd. This fact is implicit in the following development for the one-dimensional case. We wish to compute X and $Y(k)$ from $x(j)$ and $y(j)$. This is expressed by the inverse transforms

$$x(j) = \sum_{k=0}^{M-1} X(k)W_M^{jk} \quad (10)$$

$$y(j) = \sum_{k=0}^{M-1} Y(k)W_M^{jk}, \quad W_M = e^{2\pi\sqrt{-1}j/M}. \quad (11)$$

One real data set (x) is placed in the real part of the input data array and the other (y) is placed in the imaginary part so that a complex array is formed

$$\xi(j) = x(j) + iy(j). \quad (12)$$

The ξ array is then transformed to the complex frequency domain forming $Z(k)$

$$\xi(j) = \sum_{k=0}^{M-1} Z(k)W_M^{jk}. \quad (13)$$

To get the X and Y transform from the transform Z , the following development is used: multiply the y -transform expression (11) by $i = \sqrt{-1}$ and then add to and subtract it from the x -transform expression (10) which produces

$$x(j) \pm iy(j) = \sum_{k=0}^{M-1} (X(k) \pm iY(k))W_M^{jk}. \quad (14)$$

The complex conjugate of (13) can be written in terms of an inverse transform by setting $k' = M - k$ as follows:

$$\xi^*(j) = \sum_{k=0}^{M-1} Z^*(k)W_M^{-jk} = \sum_{k'=0}^{M-1} Z^*(M - k')W_M^{jk'}. \quad (15)$$

Equating coefficients of (13) with those of (14) having the plus sign (+) gives

$$Z(k) = X(k) + iY(k) \quad (16)$$

and equating coefficients of (15) with those of (14) having a minus sign (-) gives

$$Z^*(M - k) = X(k) - iY(k). \quad (17)$$

Solving these two expressions for X and Y gives

$$X(k) = \frac{1}{2}(Z(k) + Z^*(M - k)) \quad (18)$$

$$Y(k) = \frac{1}{2i}(Z(k) - Z^*(M - k)). \quad (19)$$

Thus the X and Y transforms can be resolved from the transform of the complex combination by applying the preceding expressions. This is implemented in the correlation program thereby cutting the core requirement for complex array storage in half. For the 32 by 32-

TWO DIMENSIONAL CROSS CORRELATION FUNCTION

```

A ARRAY                                B ARRAY
RUN 69002700                            RUN 69002700
LINES 79 TO 118                          LINES 89 TO 108
COLS 130 TO 169                          COLS 140 TO 159
CHAN 2                                    CHAN 6

                                RANGE= 20
                                DELTA= 6

```

```

4 4 4 4 4 3 4 3 3 4 3 3 4
3 4 4 4 4 3 4 3 3 4 3 3 4
3 4 4 5 5 4 5 4 4 4 4 3 3
2 4 5 5 5 5 5 5 4 5 4 3 2
2 3 4 5 5 5 7 6 5 5 4 3 2
2 4 4 5 6 6 8 7 6 5 4 3 2
0 1 2 3 4 6 9 7 4 4 2 1 0
0 1 2 3 4 5 7 6 4 3 2 1 0
0 1 2 3 3 3 5 4 2 3 2 0 0
0 1 2 3 3 3 4 3 2 3 2 1 0
1 2 3 4 3 2 3 2 2 3 2 1 1
2 2 2 2 1 0 1 1 1 2 2 2 2
2 3 3 2 1 1 2 1 1 2 2 2 2

```

```

MAXIMUM CORRELATION= 0.65
MINIMUM CORRELATION= 0.02
PICTURE DISTANCE AT MAX = D=( 0, 0)
EUCLIDIAN DISTANCE AT MAX= 0.0
PICTURE DISTANCE AT MIN = D=( -2, 6)
EUCLIDIAN DISTANCE AT MIN= 6.32

```

Fig. 7. Example of computer printout of 2-dimensional correlation function.

point array being used this is a saving of $2 \cdot 32 \cdot 32 = 2048$ words or 8192 bytes of storage.

The FFT method described previously significantly reduces the time required to compute the correlation function compared to the numerical integration approach. Table I presents a comparison of the time required to compute the correlation function by the two methods. The numerical integration time refers to the conventional method of computing the correlation function. The FFT time shown is the time to compute the averages and sum of squares of the two data arrays, normalize, set up the complex array, take a forward and a reverse two-dimensional Fourier transform, unscramble the two transforms as required by the core saving method discussed previously, and to extract the valid correlation function points from the total correlation function. The time savings using the FFT averages about an order of magnitude and this savings has a great impact on the usefulness of digital registration methods.

IV. ENHANCEMENT TECHNIQUES

Imagery enhancement algorithms were experimented with in an attempt to improve registration accuracy and possibly in some cases to enable registration in areas where correlation of the unaltered imagery would be impossible. Two enhancement processes are presently under study. One implements the spatial gradient edge

TABLE I
TIME COMPARISON OF NUMERICAL INTEGRATION AND FFT METHODS OF COMPUTING CORRELATION FUNCTION

Correlation	Maximum Shift in Both Directions							
	± 5 points		± 10 points		± 15 points		± 20 points	
	Area (points square)	Numerical Integration (seconds)	FFT (seconds)	Numerical Integration (seconds)	FFT (seconds)	Numerical Integration (seconds)	FFT (seconds)	Numerical Integration (seconds)
4	1.5	0.70	5.4	3.2	11.7	13.9	20.4	13.6
8	5.7	3.4	20.4	3.2	44.5	13.9	77.8	13.6
12	12.6	3.4	45.6	3.2	99.2	13.9	173.3	13.6
16	22.4	3.4	80.8	14.4	175.6	13.9	306.9	13.6
20	34.9	3.4	125.9	14.4	273.8	13.9	478.4	13.6
24	50.1	15.1	181.1	15.5	393.7	13.9	688.1	13.6

enhancement operation which amplifies the borders of features in the imagery. The other uses clustering techniques to identify homogeneous regions in imagery. It defines borders as the edges of these regions. Both schemes are based on the assumption that borders will be invariant features in imagery from different wavelength bands, whereas the reflectance and emissivity of the material within the boundaries generally varies widely across the spectrum. Also, for imagery gathered at different times the nature of the surface features may vary; however, the borders of the features tend to remain the same.

The spatial gradient is a form of the two-dimensional first derivative of a surface [6]. For continuous functions

$$|\text{grad } F| = |\Delta F| = \left[\left(\frac{\partial F}{\partial x} \right)^2 + \left(\frac{\partial F}{\partial y} \right)^2 \right]^{1/2} \quad (20)$$

In the discrete domain the first derivative is analogous to the first difference and the gradient is

$$|\Delta F_{ij}| = [(F_{i+1,j} - F_{ij})^2 + (F_{i,j+1} - F_{ij})^2]^{1/2} \quad (21)$$

A modified form of the gradient is used in the enhancement processor which computes the absolute value of the difference rather than the square. The results are similar, and computation time is saved. The multichannel gradient can also be computed by the processor. If several channels are known to be in registration the border representation can be summed to further improve the boundary estimation. The final gradient representation is thresholded to produce a binary border map which is correlated with the border map for the second image. An example of thresholded spatial gradient enhancement is shown in Fig. 8(b) for typical agricultural land. This form of enhancement is inherently "noisy" since the differentiation is in effect a high-pass filtering of the imagery. The method is also very sensitive to the border threshold used.

The clustering technique [7] is more stable but also more costly enhancement method. The technique can be applied to a single or multichannel image. An example of cluster-derived boundaries for a section of an agricultural test site is shown in Fig. 8(c). Comparison of the cluster-derived boundaries to the gradient bound-

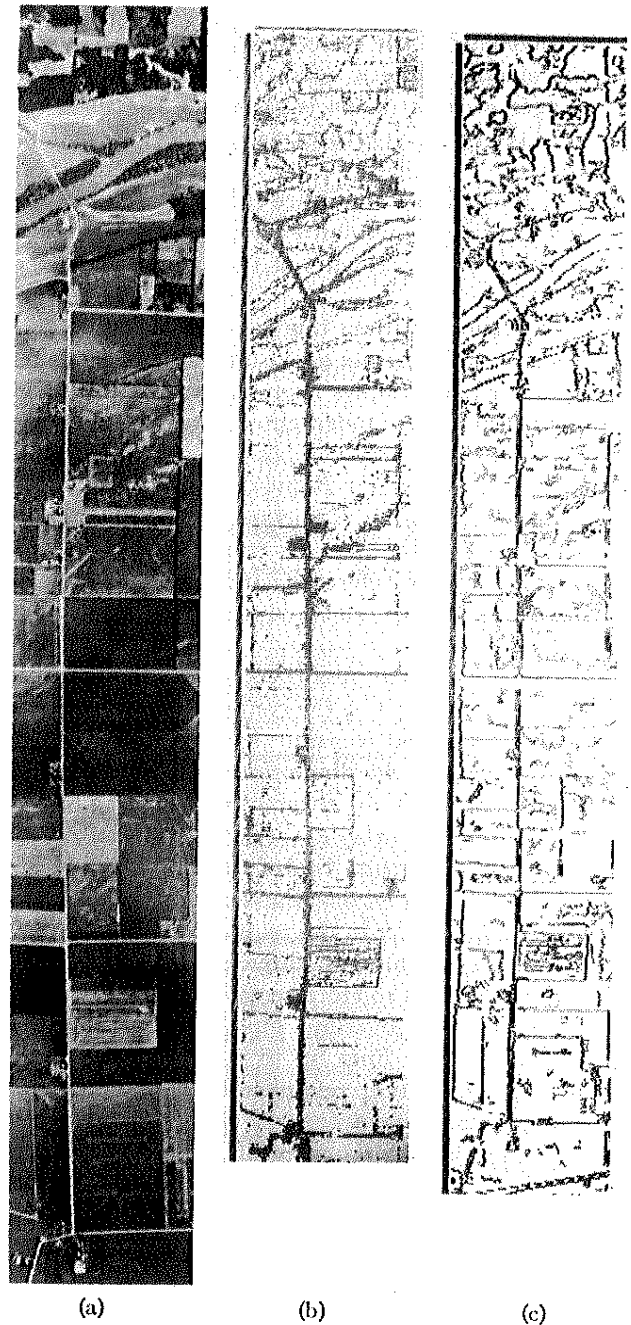


Fig. 8. Comparison of gradient and cluster-derived borders for agricultural land. (a) Aerial photo. (b) Gradient borders. (c) Cluster borders.

aries shows improvement in uniformity and detail using the clustering technique. Either method of enhancement can be executed for the arrays being correlated in the registration system. Some results of correlation of border enhanced imagery are presented in Section VI.

V. REGISTRATION SYSTEM

The enhancement and correlation algorithms form the core of the imagery registration system. The overlay or image transformation function, control, and input-output functions complete the software system which operates via Disk Programming System 44 on an IBM System 360 Model 44 computer. A general block diagram of the system is presented in Fig. 9. The imagery to be registered is referred to in terms of apertures (derived from the concept of a segment or window of the electromagnetic spectrum). A set of one or more images in registration is called an aperture. There can be three input apertures denoted, *A*, *B*, and *C*. The *A* aperture is called the context or reference aperture. The *B* and *C* apertures are called overlay apertures and are correlated and shifted with respect to the context aperture. The segments of the system are discussed as follows.

- 1) **Input monitor:** this routine selects imagery lines from the appropriate input file and extracts the desired channels and samples from the line. A fully flexible input definition is implemented since the three apertures could exist in many combinations, i.e., context on tape *A*, overlay *B* on tape *B*, overlay *C* on tape *C*, or context on tape *A*, overlay *B* on tape *A*, overlay *C* on tape *B*, etc. This feature allows a context aperture to be used for a geometric reference, and all new imagery of the same area can be overlaid such that the same geometric dimensions are achieved for all the imagery processed. Two image sets (overlay *B* and *C*) can be overlaid with respect to the context image (*A*) in one pass.

- 2) **Scaling and rotation:** scale factors are applied to the imagery in two dimensions. The input monitor and scaling operation are interactive since the line required by the *B* or *C* overlay aperture depends on the scale factor for that aperture. Global rotation is implemented by reading the desired line segments from tape and placing them in the buffer in the required position.

- 3) **Aperture buffers:** three large cyclic or processing buffers are used to store rectangular imagery arrays for correlation and overlay. This approach allows addressing of imagery lines up to several lines ahead of and behind the current reference line without tape input. Rectangular arrays for correlation can be extracted from any area in the span of available lines, and final rotation and scaling are implemented by addressing the appropriate image cell in the overlay aperture buffers.

- 4) **Enhancement processor:** the spatial gradient and cluster border-finding algorithm are presently implemented for enhancement. Also, an image complexity estimator is included with the gradient processor which

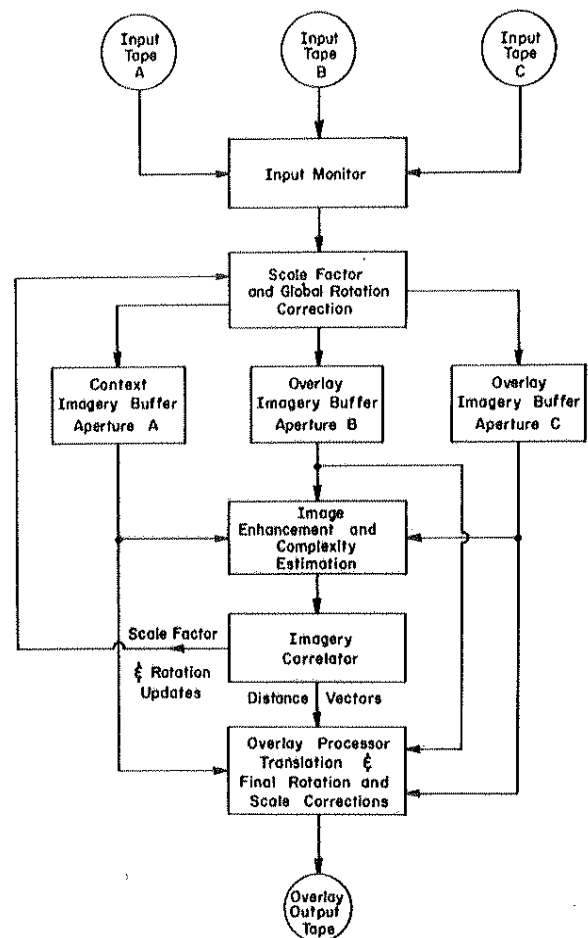


Fig. 9. Digital imagery registration system.

supplies an index of the correlatability of two image arrays [8]. If the index is too low or too high, correlation is not attempted since lock-on failure is highly likely in this case. The complexity index can also be used to vary the correlation integration range; however, the FFT correlation speed is not greatly improved as was the case for numerical integration method discussed in [8] since the total correlation time is largely input-output time. Other methods of enhancement are being studied such as frequency-domain filtering to achieve border enhancement, and work is continuing on improvement of the clustering algorithm for border finding.

- 5) **Correlator:** the FFT correlation process is implemented for square arrays up to 32 by 32 points. The correlation array size minus the integration range is twice the maximum shift in both dimensions. An integration range of 20 and an array size of 32 is most commonly used, and thus the maximum shift is ± 6 points. A numerical integration correlator is also implemented which can search over a larger range to find an initial lock-on point. The FFT correlator can then determine the point of maximum correlation for changes of up to 6 points between points of correlation. The point of maximum positive and the point of maximum negative

correlation is found, and the distance coordinates and euclidian distance are computed for each.

Correlation is carried out at the intersections of a grid as pictured in Fig. 10. The line and column interval between correlation points depends on the amount of variation in the misregistration between two apertures. In some cases, such as aircraft scanner data in which random distortion exists, correlation intervals of 25 lines and columns have been used. For scanned aerial and satellite photographs taken from the same position, the variation in misregistration is relatively small, and intervals of 100 points are used. Also, an averaging of misregistration estimations is achieved by closely spaced correlations as long as the computation time does not become excessive. The system can operate in an adaptive mode by employing either of two correlation estimators. The enhancement processor computes the average number of border points per unit area, and previous work has shown that if this estimator is above or below certain limits correlation is likely to fail. Secondly, if the absolute value of the maximum or minimum correlation is below a threshold it is also very likely that the lock-on point is badly in error. Thus the correlation output is discarded if the estimator indicates erroneous results are likely.

A correlation function interpolator is included which uses the value of the peak and its eight or 24 neighbors to estimate the true point of maximum or minimum correlation to better than one resolution element.

6) Overlay: the line and column registration errors from the correlator are used in a linear least-squares procedure which determines a two-dimensional shift function defined over the entire space between line and column correlation nodes shown in Fig. 10. The shift function is of the form

$$\begin{aligned} \Delta L &= a_{L1}L + a_{L2}C + a_{L3} \\ \Delta C &= a_{C1}L + a_{C2}C + a_{C3} \end{aligned} \quad (22)$$

where

- ΔL line shift value
- ΔC column shift value
- L line
- C column
- a_L, c coefficients determined by a least-squares fit to correlation distances.

This function implements translation in two dimensions, small rotations and scaling, depending on the values of the coefficients. For example, if the first and second coefficients are 0, the function is a translation of a_{L3} lines and a_{C3} columns. If a_{L2} and a_{C1} are 0, then the result is scaling and translation in two dimensions. If $a_{L1} = 1 - \cos \theta$, $a_{L2} = -\sin \theta$, $a_{C1} = \sin \theta$, and $a_{C2} = 1 - \cos \theta$, the function is a rotation through an angle θ about the point $L=0, C=0$. The linear shift function is updated each line where correlation takes place. Also, the scale factor and global rotation is determined at each update

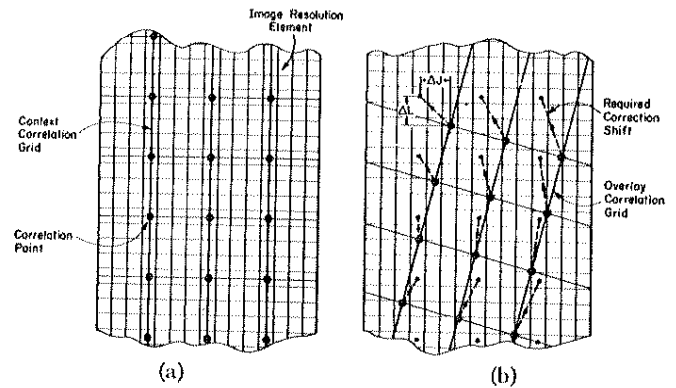


Fig. 10. Context and overlay aperture correlation and correction grids. (a) Context aperture. (b) Overlay aperture.

and is fed back to the input monitor. Translation corrections are also fed back so that the amount of correction that the shift function must handle is minimal.

Output lines are assembled in a core buffer by addressing appropriate areas in the imagery stored in core. The overlay tape format is one line per block with a maximum number of 30 channels. The total record byte length is the product of the number of channels and the samples per line and is limited only by the core buffer size available to read and write the line. All imagery is represented in byte form (8-bit words) on tape and is also processed in the computer in byte form. The processing rate of the present system is relatively slow due to nonoptimum input-output and computational procedures. The correlation time averages 3.5 seconds, gradient enhancement of two 32 by 32 arrays requires 2 seconds, and cluster border enhancement of the same arrays requires 10 seconds. The input, overlay, and output rate exclusive of enhancement averages about 1 second per line of 2800 samples. The processing rate has not been one of the primary considerations during development of the system. Once a satisfactory set of algorithms has been defined, program refinement can be carried out to optimize the processing speed.

VI. CORRELATION ANALYSIS

The two dimensional FFT array correlator portion of the system was used to analyze real imagery to determine the relationship between the correlation coefficient for imagery from different bands and the spatial registration accuracy. Imagery known to be in registration from 17 bands listed in Table II was correlated over a test area to obtain basic trends and nominal values for the accuracy that the registration system can achieve for a specific example.

Correlation ranges from four to over 20 image points square were used in tests to determine the required range for consistent lock-on of the correlator. A correlation range of 20 points square tended to give satisfactory performance for the rectangular agricultural field type of ground cover encountered in the data analyzed [8].

TABLE II
SPECTRAL BANDS STUDIED

Channel Number	Aircraft Scanner Wavelength Bands		Description
	Channel	Wavelength Limits (μm)	
1		0.40-0.44	violet
2		0.46-0.48	blue
3		0.52-0.55	green
4		0.55-0.58	green
5		0.58-0.62	yellow
6		0.62-0.66	red-orange
7		0.66-0.72	red
8		0.72-0.80	reflective infrared
9		0.8-1.0	reflective infrared
10		1.0-1.4	reflective infrared
11		1.5-1.8	reflective infrared
12		2.0-2.6	reflective infrared
13		4.5-5.5	thermal infrared
14		8-13.5	thermal infrared
Apollo 9 Multispectral Photography Bands			
1		0.47-0.61	blue-green
2		0.59-0.715	red
3		0.68-0.89	reflective infrared

The correlation coefficient as a function of wavelength for several individual cover types and for a large area (global case) is presented in Fig. 11(a) and (b). The blue wavelength band (0.46 to 0.48 μm) was used as reference, and data in this band was correlated with all other bands available. The cover types, corn, soybeans, wheat, and bare soil, chosen from June, 1969, aircraft scanner data were used. The June data is agriculturally diverse in that many different crops are present in various stages of growth. Three dips in the correlation curve are seen at about 0.56, 0.9, and above 5 μm with a reversal centered at 0.9 μm for two cover types. The general shape of the curves is similar for the individual cover types and for the global or random sample case (Fig. 11(b)). It was hypothesized that the higher the correlation the higher will be the registration accuracy. Also, it seemed that for strongly negatively correlated regions, the negative peak of the cross correlation function should be used as the point of lock-on.

The distance of the cross correlation function peak from the origin (Euclidian distance) was then plotted and is shown in Fig. 11(c). It can be observed that the error generally follows the expected trend in which high negative or positive correlation reduces the error. Increasing distance in wavelength between the two channels being correlated consistently increases the error so the effects due to correlation error are superimposed on the general increase from 0.47 to 11 μm . A similar analysis was run for three band-digitized Apollo 9 multispectral photography from Imperial Valley, Calif. The results are presented in Fig. 12(a) and (b). The inverse relationship of correlation to registration accuracy is clearly seen here. The error magnitudes are considered to be reasonable, but in many applications even an average error magnitude of one or two resolution elements

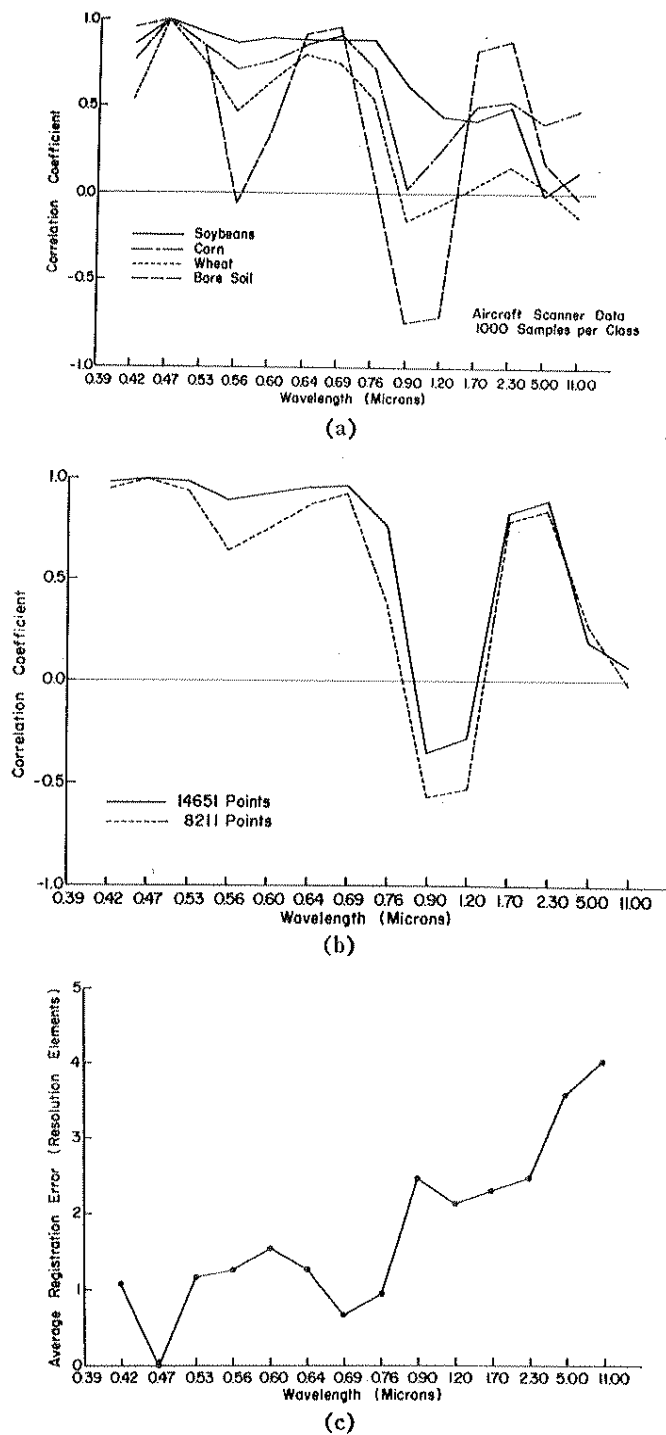


Fig. 11. (a) Correlation of blue (0.46-0.48 μm) channel with all other channels for agricultural features in June, 1969 (test site PF24). (b) Global correlation of blue (0.46-0.48 μm) channel with all other channels for agricultural test area PF24 in June, 1969. (c) Average registration error for correlation of blue (0.46-0.48 μm) channel with all other channels for agricultural land in June, 1969 (test site PF24).

is not acceptable. The results indicate that the farther apart the bands are the greater the error and, that even adjacent channels do not achieve perfect registration. This result may be pessimistic since the sample was small (30 correlations over 111 000 image points) and for agricultural land only. Further correla-

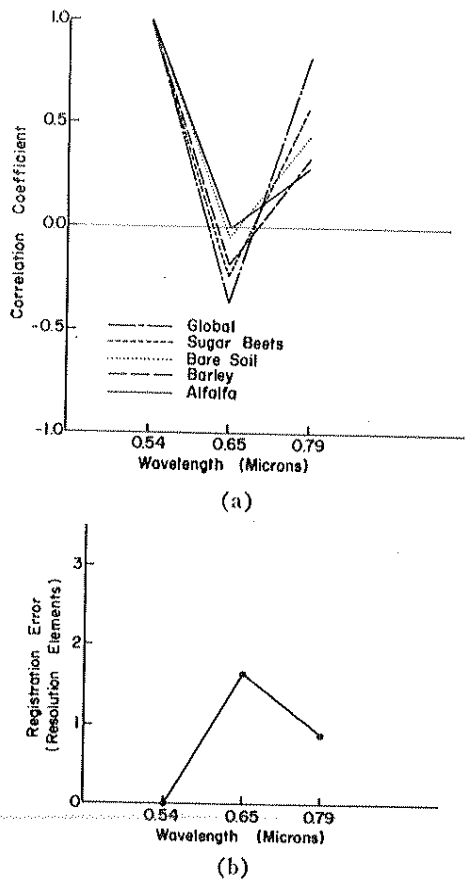


Fig. 12. (a) Correlation of green channel (0.47-0.61 μm) with red and infrared for Apollo 9 multispectral photography (SO65 test site 15A). (b) Average registration error for correlation of green band (0.47-0.61 μm) with red and infrared from Apollo 9 photography.

tion tests were run to see if an improvement could be achieved using enhanced imagery.

Test imagery enhanced by the spatial gradient method was correlated for the test area, and the error curve is plotted in Fig. 13 along with the unenhanced error curve from Fig. 11(c). The performance for this test was not significantly better than for the unaltered data case. The clustering technique was then employed for boundary enhancement, and the correlation tests were rerun using the results. The plotted error curve is also shown in Fig. 13. It can be seen that the performance is about the same as for the unenhanced case with some improvement at 0.53 and 0.60 μm and substantial improvement above 2.3 μm .

The correlation analysis presented here is intended to illustrate the worst situation in which the shortest wavelength channel is used as reference and is correlated with other channels as distant as 8 to 14 μm . In general, the optimum approach is to use the two most highly correlated channels from the misregistered image sets as input to the correlator. Often the same wavelength band may be available in the two images to be registered, and in this case highly accurate registration may be possible. The results here indicate that correlation of unaltered imagery is about as good an approach

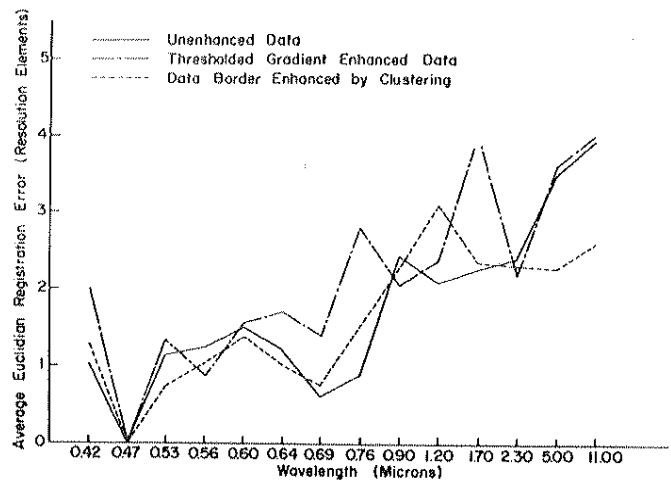


Fig. 13. Average registration error for correlation of blue (0.46-0.48 μm) channel with all other channels for agricultural test site PF24 in June, 1969.

as using costly enhancement before correlation, and in some cases it is better. If accurate correlation of widely separated bands is required, the cluster boundary enhancement approach appears promising.

VII. APPLICATIONS

Multispectral and multitemporal imagery can be used as input to statistical pattern-recognition systems for automatic classification of scene features [1]. In such a process several spectral measurements from one image resolution element are used to classify that element into one of a set of scene feature classes. This process requires that all the spectral measurements come from the same resolution element, and it is the purpose of the registration system to achieve this state. Once registered imagery is obtained, the pattern-classification algorithms can be applied to any recognition task. Two examples of multispectral pattern classification made possible by image registration are shown in Figs. 14 and 15. In example 1, aircraft scanner data from 1.0-1.4-, 1.5-1.8-, 2.0-2.6- μm bands were registered with multi-band data from the visible and near-reflective infrared portion of the spectrum from 0.4 to 1.0 μm . A feature selection algorithm [9] was employed to determine the most distinctive four bands for separation of the pattern classes considered, corn, soybeans, wheat, oats, pasture, grass, trees, and water, from data collected in June, 1969. The four bands chosen by the feature selection algorithm were 0.4-0.44, 0.66-0.72, 1.0-1.4, and 1.5-1.8 μm , and 290 290 points were classified for agricultural test site PF24 using the LARS multispectral pattern-recognition system. In Fig. 14 the classifier output is coded to denote the classifier decisions. The average classification accuracy based on 87 679 test points was 80 percent. Classification accuracy using only the best four bands available before registration was 60 percent. Use of the two infrared channels above 1.0 μm was made possible by the registration process.

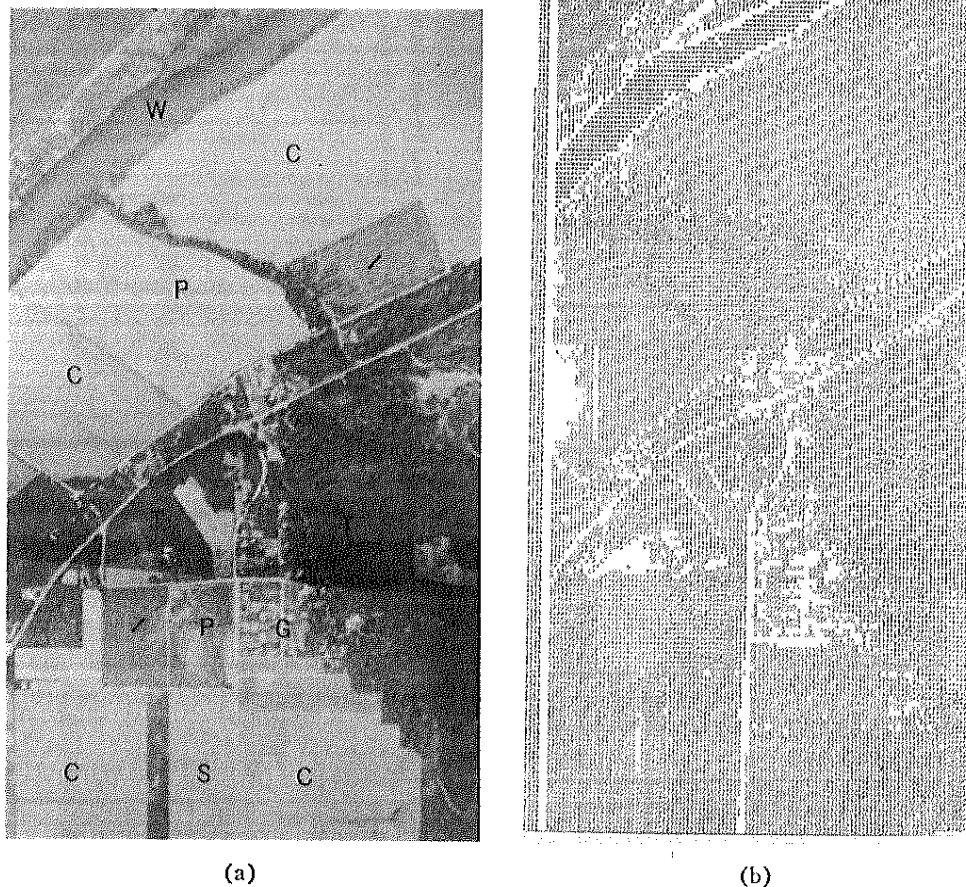


Fig. 14. Computer classification of agricultural test site PF24 using registered multispectral scanner imagery. (a) Aerial photograph. (b) Classification printout. Key: C, corn; S, soybeans; /, wheat; O, oats; P, pasture; G, grass; T, trees; W, water.

In the second experiment the digitized multiband photography from Apollo 9 taken over Imperial Valley, Calif., was registered, and statistical pattern classification was attempted. The available wavelength bands were listed in Table II. The basic pattern classes, green vegetation, bare soil, salt flat, and water, were defined, and classification was carried out on 66 000 samples in the NASA S065 test site near Brawley, Calif. Fig. 15 is a printout of the results which proved to be quite accurate. The average performance was 95 percent for 880 test samples. The accuracy using only one channel was only 80 percent. Here again the multispectral registration system enabled use of automatic statistical pattern-recognition techniques for classification of imagery.

Multispectral classification has been carried out at LARS on registered imagery from agricultural, geological, and hydrological test sites. Imagery obtained at different times offers a valuable dimension, and registration of multitemporal imagery is in progress. Analysis will begin when data from several time points are available for a test site. The LARS imagery registration system is a key part of the remote-sensing research project and it is making valuable new dimensions available to the researcher.

Multi-images to be used for multidimensional analysis should ideally be obtained in registered form so that costly and, at best, inexact registration processing is not necessary. At the present time perfectly registered multispectral and multitemporal imagery is not available to all who require it. The software system described here is one approach to achieving registration where the need exists. It is suitable for a research environment where a small quantity of imagery is needed for design of algorithms and data-handling systems. Specialized high-speed image-registration and data-handling systems can then be built based on results of the research using this slower general-purpose implementation. The LARS registration software system can be operated on any general-purpose digital computer having FORTRAN IV capability and sufficient memory capacity. The largest phase of the current system requires 128 000 bytes of storage.

VIII. CONCLUSIONS

A software system for spatial registration of multiple images has been described, and some analysis results were presented to indicate likely system performance. The use of the FFT for improving the speed of calcula-

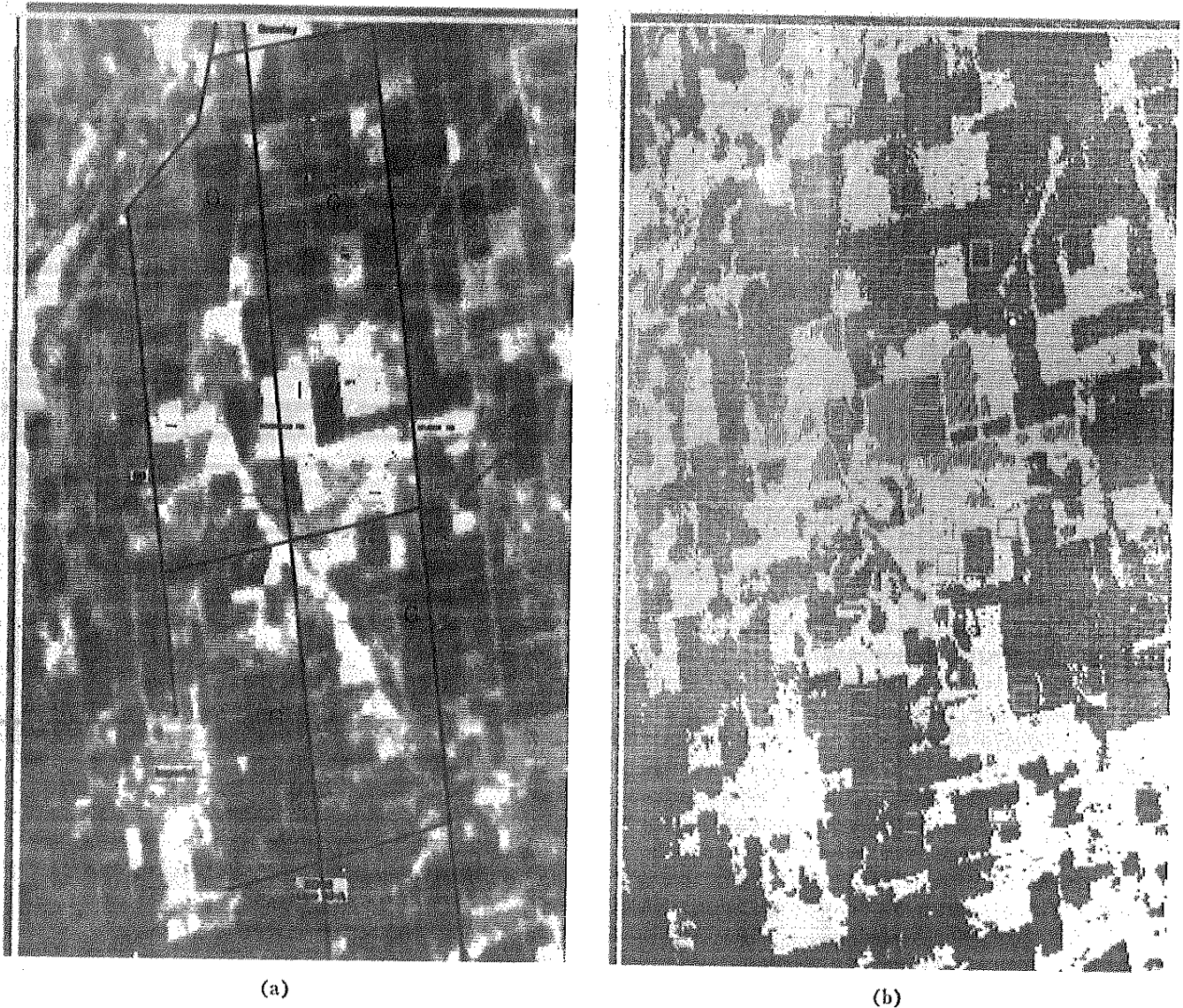


Fig. 15. Computer classification of surface features in Apollo 9 photography of S-065 test site. (a) Gray-scale printout. (b) Classification printout. Key: G, green vegetation; —, soil; I, salt flats; W, water.

tion of the image correlation function was discussed in detail. Two methods of imagery enhancement were described, and test results were presented. The structure of the system was outlined, and utilization of registered imagery produced by the system was described.

The goal of the registration system development project at LARS is to achieve precise registration of digital multi-image data so that research could be performed by automatic multivariate analysis techniques for imaging remote-sensor data. The current system is producing registered imagery from an aircraft multispectral scanner and from digitized aerial and space photography. This imagery is being used at LARS for automatic classification research of earth-surface features in the categories of agriculture, geology, hydrology, transportation, forestry, and detection of pollution. Imagery collected over an area at different times is

being registered to enable study of the time-varying properties of surface features. The registration system is thus enabling advanced research to be conducted into the automation of earth-resources survey processes vital to the preservation and improvement of life on the earth.

ERTS will produce imagery over large areas of the earth's surface in three bands in photographic and digital form. Each image frame will cover approximately a 100 mi² and will contain several million picture points in each band. Digitization if necessary and registration of this imagery over areas of interest on a frequent time base will enable detailed study of the time-varying properties of these surface features. It is predicted that ERTS will photograph every point on the earth every 20 days; thus a time history of over 18 samples per year will be available for each point. Also,

registration of imagery from other sensors in other parts of the spectrum will enable multidimensional analysis similar to that made possible by wideband line scanners. Other applications for registered multi-images will arise as research progresses. The LARS system described here is designed to handle a wide range of requirements for research on limited quantities of data.

ACKNOWLEDGMENT

The airborne multispectral scanner data used in the analysis were gathered by the Institute of Science and Technology, University of Michigan, Ann Arbor, under contract to NASA using equipment supplied by the U. S. Army Electronics Command. The Apollo 9 satellite photography was obtained from the Earth Observations Division, Manned Spaceflight Center, NASA, Houston, Tex., and was digitized by Optronics International, Chelmsford, Mass.

REFERENCES

- [1] K. S. Fu, D. A. Landgrebe, and T. L. Phillips, "Information processing or remotely sensed agricultural data," *Proc. IEEE*, vol. 57, pp. 639-653, April 1969.
- [2] D. Steiner, "Time dimension for crop surveys from space," *Photogrammetr. Eng.*, vol. 36, pp. 187-194, February 1970.
- [3] B. Ostle, *Statistics in Research*. Ames, Iowa: Iowa State University Press, 1963.
- [4] J. W. Cooley and J. W. Tukey, "An algorithm for the machine calculation of complex Fourier series," *Math. Comput.*, vol. 19, pp. 297-301, April 1965.
- [5] J. W. Cooley, P. A. Lewis, and P. D. Welch, "The fast Fourier transform algorithm and its applications," IBM Watson Research Center, Yorktown Heights, N. Y., IBM Res. Rep. RC1743, February 9, 1965.
- [6] A. Rosenfeld, *Picture Processing by Computer*. New York: Academic Press, 1969.
- [7] A. G. Wacker, "A cluster approach to finding spatial boundaries in multispectral imagery," Laboratory for Applications of Remote Sensing, Purdue University, West Lafayette, Ind., LARS Information Note 122969.
- [8] P. E. Anuta, "Digital registration of multispectral video imagery," *SPIE J.*, vol. 7, pp. 168-175, September 1969.
- [9] P. J. Min and K. S. Fu, "On feature selection in multiclass pattern recognition," School of Elec. Eng., Purdue University, Lafayette, Ind., Tech. Rep. TR-EE68-17, July 1968.