F. Ortí
IBM Scientific Center
Madrid, Spain

Optimal Distribution of Control Points to Minimize Landsat Image Registration Errors

Using as an optimality criterion the mean square registration error for the whole image, the optimum location of ground control points was found to be around certain locations on the left and right edges of the Landsat-MSS image.

Introduction

To transform a digital image of the Earth's surface, taken by a satellite sensor, into a corrected image according to some cartographic projection, one needs to know the mathematical functions relating the coordinates of the two images (raw and corrected), and to have a rule to assign intensities to the pixels of the corrected image.

For MSS images the problem of obtaining the geometric transformation functions is essentially that of finding the intersection of the scanner beam with the Earth's surface as a function of time, as time is easily related to the raw image coordinates. An accurate knowledge of the Earth's geometry, the trajectory of the satellite, its attitude, and the position of the scanner beam with respect to the satellite are necessary in order to derive the mapping functions relating the raw and corrected images. Once these functions have been determined, and a choice has been made regarding the resampling technique to assign the corresponding intensity values to the pixels of the corrected image, the actual mapping can be carried out (see Bernstein, 1976).

Abstract: *To precisely correct an MSS-Landsat image, ground control points are necessary because of the inaccuracy of the satellite's attitude and altitude measurements. If the attitude and altitude are assumed to be described by certain polynomials of time, the corresponding coefficients can be estimated from the set of GCP's, and their estimated error propagated to obtain an average registration error over the whole image as a function of the GCP's coordinates. Minimization of this error leads to the result that GCP's should be chosen around certain locations on the left and right edges of the image. Some experiments are run to assess the practical value of this result.*

For Landsat images all necessary items to obtain the mapping functions will be assumed to be known with sufficient accuracy, with the exception of the attitude and altitude of the satellite which have to be estimated from a set of ground control points (GCP). Furthermore, following Bernstein (1973), it will be assumed that within a Landsat frame the three attitude angles can be adequately described by cubic polynomials of time, and the altitude by a linear one, reducing the estimation problem to one of determining the corresponding 14 coefficients from the given set of GCP's.

Obviously, the error in the determination of these coefficients, and thus the registration error, will be dependent upon the number of GCP's, their location error, and their spatial distribution. The last item is the subject of this paper, and although the analysis is carried out for the attitude and altitude models stated above, other models can be treated along the same lines.

Photogrammetric Engineering and Remote Sensing,
Vol. 47, No. 1, January 1981, pp. 101-110.

Mean Square Registration Error

In what follows only the registration error due to the uncertainty in location of the GCP's will be considered (Earth geometry, satellite velocity, etc., are assumed to be perfectly known). This uncertainty will cause an error in the values of the computed coefficients of the attitude-altitude model, and thus an error in the position of each point to be transformed by the mapping functions. The average of this error taken over all points in the image will serve as a criterion for the goodness of registration of the image to the corresponding cartographic projection. The aim of this paper is to find the optimal spatial distribution of GCP's which will minimize this average registration error.

A brief summary of the least-squares method used to obtain the model coefficients from the GCP's seems in order to derive the expression for the registration error of a point in the image.

Figure 1 shows the geometry of the problem, and by inspecting it one can derive the expressions for the differences in position for a GCP in terms of the attitude and altitude deviations about their nominal values.

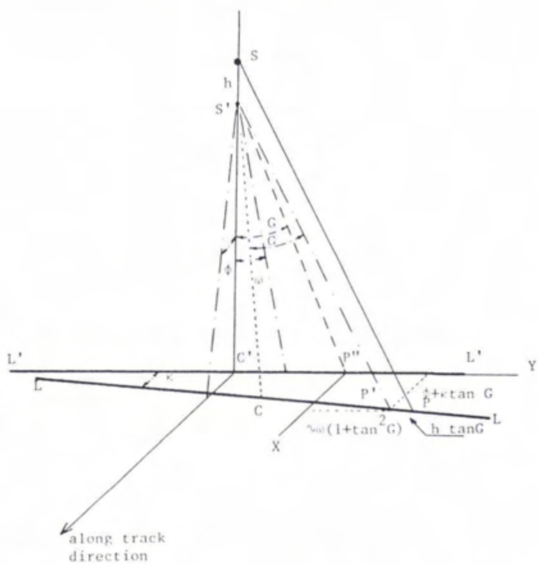

Fig. 1. Geometry of the problem: S is the actual position of the scanner (S' is the nominal); LL is an actual scan line ($L'L$ is the nominal position; P is the actual position of a point, P' is the position for nominal altitude, and P' is the position for nominal attitude and altitude. (All angles have been made much larger than they actually are for Landsat. The nominal altitude $S'C'$ has been taken as unit length.)

Assuming these deviations to be small, and measuring lengths in terms of the nominal altitude, the following dimensionless equations result:

$$x = \phi + \kappa \tan G(p) \tag{1}$$

$$y = \omega\,[1 + \tan^2 G(p)] + h \tan G(p) \tag{2}$$

where p is the GCP's pixel number within the scan line measured from its center; $G(p)$ is the corresponding scan angle, zero at the center of the scan line; ϕ, ω, and κ are the attitude angles, pitch, roll, and yaw, and h is the relative altitude deviation, all corresponding to the instant in which the GCP was seen by the MSS; and x and y are the differences in position of the GCP along and across the orbital track on a plane tangent to the Earth through the nadir point of the image, and are computed as differences between the position of the GCP obtained from the given map coordinates and the position of the GCP obtained from transforming its raw image coordinates without consideration of any attitude or altitude deviations from the nominal values (x and y are made dimensionless with the nominal altitude).

Equations 1 and 2 show that pitch and yaw are decoupled from roll and altitude, and therefore the corresponding estimation problems can be treated independently.

It has been previously stated that the attitude angles and the altitude of the satellite are assumed to be well presented by cubic and linear polynomials of time; time and scan line number are linearly related if the scanning time is considered negligibly short; therefore, the attitude angles and the altitude may be considered as cubic and linear polynomials of the scan line number or of any linear function of it. In what follows a normalized line number coordinate will be used, taking the values -1 and 1 at the top and bottom of the image respectively. One may then write:

$$\phi = \sum_{i=0}^{3} \phi_i P_i(l) \qquad \kappa = \sum_{i=0}^{3} \kappa_i P_i(l)$$
$$\omega = \sum_{i=0}^{3} \omega_i P_i(l) \qquad h = \sum_{i=0}^{1} h_i P_i(l) \tag{3}$$

where l is the normalized line scan number in the interval $(-1, 1)$; ϕ_i, ω_i, κ_i, and h_j $(i=0 \ldots 3, j=1)$ are the 14 unknown coefficients to be determined; and $P_i(l)$ are the orthonormal Legendre polynomials of i^{th} degree (see, for example, Courant and Hilbert (1953)).

The reason for using the Legendre polynomials instead of the successive powers of l is to simplify the expression for the mean square registration error when integrating over the image.

Substituting the expressions in Equations 3 into Equations 1 and 2, calling $F = \tan G(p)$, and introducing the vectors $\boldsymbol{v}^T(l) = [P_0(l), P_2(l), P_1(l), P_3(l)]$, $\boldsymbol{u}^T(l) = [P_0(l), P_1(l)]$, $\boldsymbol{c}_x^T = (\phi_0 \ldots \phi_3, \kappa_0 \ldots \kappa_3)$, and $\boldsymbol{c}_y^T = (\omega_0 \ldots \omega_3, h_0, h_1)$, results in

$$x = [\boldsymbol{v}^T(l), F.\boldsymbol{v}^T(l)] .\boldsymbol{c}_x$$
$$y = [(1 + F^2)\boldsymbol{v}^T(l), F\, \boldsymbol{u}^T(l)].\boldsymbol{c}_y \tag{4}$$

Writing Equations 4 for N ground control points with raw coordinates given by the normalized line scan number l_j and $F_j = \tan G(p_j)$ $(j = 1 \ldots N)$, the following set of equations is obtained

$$x_j = [\boldsymbol{v}^T(l_j), F_j\boldsymbol{v}^T(l_j)].\, \boldsymbol{c}_x + e_{xj}$$
$$y_j = [(1 + F_j^2)\boldsymbol{v}^T(l_j), F_j\boldsymbol{u}^T(l_j)].\, \boldsymbol{c}_y + e_{yj} \tag{5}$$

where e_{xj} and e_{yj} are the location errors associated with the j^{th} GCP.

In matrix notation these equations are rewritten as

$$\mathbf{X} = \mathbf{W}_x \cdot \boldsymbol{c}_x + \mathbf{E}_x$$
$$\mathbf{Y} = \mathbf{W}_y \cdot \boldsymbol{c}_y + \mathbf{E}_y \tag{6}$$

Assuming the errors associated with different GCP's to be statistically independent, with zero mean and the same variance for all, though different in the two directions, the solution of Equations 6 in a least-mean-square sense is given by (see for example Hamilton (1964))

$$\hat{\boldsymbol{c}}_x = \frac{1}{\sigma_x^2}\, \boldsymbol{B}_x^{-1}\, \mathbf{W}_x^T\, \mathbf{X} \tag{7}$$

$$\hat{\boldsymbol{c}}_y = \frac{1}{\sigma_y^2}\, \boldsymbol{B}_y^{-1}\, \mathbf{W}_y^T \mathbf{Y} \tag{8}$$

where σ_x^2 and σ_y^2 are the variances of the location error of the GCP's,

$$\boldsymbol{B}_x = \frac{1}{\sigma_x^2}\, \mathbf{W}_x^T\, \mathbf{W}_x,$$

and

$$\boldsymbol{B}_y = \frac{1}{\sigma_x^2}\, \mathbf{W}_y^T\, \mathbf{W}_y.$$

The covariance matrices of the estimates are given by

$$\text{Cov}\,(\hat{\boldsymbol{c}}_x) = \boldsymbol{B}_x^{-1}$$
$$\text{Cov}\,(\hat{\boldsymbol{c}}_y) = \boldsymbol{B}_y^{-1} \tag{9}$$

And the variance of the propagated registration error in a point (l, F) is (cf Equation 4)

$$\mathrm{var}\,(\epsilon_x) = (\boldsymbol{v}^T, F\boldsymbol{v}^T)\,\boldsymbol{B}_x^{-1}\begin{pmatrix}\boldsymbol{v}\\ F\boldsymbol{v}\end{pmatrix} \tag{11}$$

$$\mathrm{var}\,(\epsilon_y) = [(1 + F^2).\boldsymbol{v}^T, F\boldsymbol{u}^T]\boldsymbol{B}_y^{-1}\begin{pmatrix}(1 + F^2)\,.\boldsymbol{v}\\ F\boldsymbol{u}\end{pmatrix} \tag{12}$$

The expected value of the square of the total registration error in one point is

$$\epsilon(l,F) = \mathrm{E}\,\{\epsilon_x^2 + \epsilon_y^2\} = \mathrm{var}\,(\epsilon_x) + \mathrm{var}\,(\epsilon_y) \tag{13}$$

and the average over the whole image

$$\epsilon_{lF} = \frac{1}{4F_M}\int_{-1}^{1}\int_{-F_M}^{F_M} \epsilon(l,F)dldF \tag{14}$$

where F_M is the maximum value F can take over the image (the origin is taken at the center).

The mean square registration error, ϵ_{lF}, which through $\boldsymbol{B}_x^{-1}$ and $\boldsymbol{B}_y^{-1}$ depends on the GCP's coordinates, shall be the criterion to be minimized with respect to the locations of the GCP's; the square root of this quantity will give a good indication of how well the raw image can be corrected with a given set of GCP's.

Note that, although ϵ_{lF} has been called the mean square registration error, it is computed on a plane tangent to the Earth at the nadir point of the image, not in the cartographic projection coordinate system.

Some comments on several points of this paragraph should be made:

- The use of linear expressions in the deviations from nominal attitude and altitude as written in Equations 1 and 2 might not be accurate enough for some Landsat image. This can be overcome by an iterative procedure, including in the left-hand side of the equations the corresponding non-linear terms, with a guessed set of values for the attitude and altitude coefficients. From there a new set of values is obtained through Equations 7 and 8, and the process is repeated until the differences between the values in two successive iterations are less than prespecified amounts. Due to the smallness of the non-linear terms, their effect on the optimal distributions of GCP's is expected to be negligible, though their effect on the registration error in some parts of the image could be important.
- The effect of considering one image line per scan, as is done in writing Equation 3, instead of considering six lines per scan, the actual case in Landsat image, is assumed to be negligible.
- To assume that all GCP's have the same variance for the location error is probably not far from the truth if the method followed to locate the GCP's in the raw image does not rely on finding a single point, but rather a set of them. For example, locating a GCP by optically or otherwise superimposing a map on an area around the GCP taken from the raw image (after some corrections have been applied) is believed to produce a location error which will be quite independent of the feature (road junction, dam, airport) represented by the surroundings of the GCP.

The variances of the location error, assumed to be the same for all GCP's, can be estimated after the fit by

$$\sigma_x^2 = \frac{(\mathbf{X} - \mathbf{W}_x \cdot \hat{\mathbf{c}}_x)^T(\mathbf{X} - \mathbf{W}_x \cdot \hat{\mathbf{c}}_x)}{N - 8},$$
$$\sigma_y^2 = \frac{(\mathbf{Y} - \mathbf{W}_y \cdot \hat{\mathbf{c}}_y)^T(\mathbf{Y} - \mathbf{W}_y \cdot \hat{\mathbf{c}}_y)}{N - 6} \tag{15}$$

Solution of the Optimization Problem

The problem to be solved is the following: find the coordinates of N ground control points, which minimize

$$\epsilon_{lF} = \frac{1}{4F_M}\int_{-F_M}^{F_M}\int_{-1}^{1} \epsilon(l,F)dFdl \tag{16}$$

subject to the constraints

$$|l_i| \leq 1,\ |F_i| \leq F_M, \qquad i = 1,2,\ldots, N \tag{17}$$

Referring to the Appendix, the expression for the registration error at a point, given by Equation 13 can be written as

$$\epsilon(l,F) = \sigma_x^2\,(\boldsymbol{v}^T\mathbf{X}_x\boldsymbol{v} + 2F\boldsymbol{v}^T\mathbf{Z}_x\boldsymbol{v} + F^2\boldsymbol{v}^T\mathbf{Y}_x\boldsymbol{v}) + \sigma_y^2\,[(1 + F^2)^2\boldsymbol{v}^T\mathbf{X}_y\boldsymbol{v} + F(1 + F^2)\,(\boldsymbol{u}^T\mathbf{Z}_y\boldsymbol{v} + \boldsymbol{v}^T\mathbf{Z}_y^T \cdot \boldsymbol{u}) + F^2\boldsymbol{u}^T\mathbf{Y}_y\boldsymbol{u}] \tag{18}$$

where the expressions for the matrices X_x, Y_x, Z_x, X_y, Y_y, and Z_y are written in the Appendix.

Averaging Equations 18 over the image, taking into account the orthonormality of the Legendre polynomials, results in

$$\epsilon_{lF} = \frac{\sigma_x^2}{2} \text{tr} \left[X_x + \frac{F_M^2}{3} Y_x\right] + \frac{\sigma_y^2}{2} \text{tr} \left[\left(1 + \frac{2}{3} F_M^2 + \frac{1}{5} F_M^4\right)X_y + \frac{F_M^2}{3} Y_y\right] \tag{19}$$

where tr is the trace operator.

Substituting the expressions for the matrices X_x, Y_x, X_y, Y_y given by Equation A.6 of the Appendix, Equation 19 may be rewritten as

$$\begin{aligned}\epsilon_{lF} = {} & \frac{\sigma_x^2}{2} \ \text{tr} \left[M_x^{-1} + M_x^{-1}Q_x^T Y_x Q_x M_x^{-1} + \frac{F_M^2}{3} \ (R_x^{-1} + R_x^{-1}Q_x X_x Q_x^T R_x^{-1})\right] \\ & + \frac{\sigma_y^2}{2} \text{tr}\left[\left(1 + \frac{2}{3} F_M^2 + \frac{1}{5} F_M^4\right) (M_y^{-1} + M_y^{-1}Q_y^T Y_y Q_y M_y^{-1}) \right. \\ & \left. + \frac{F_M^2}{3} \ (R_y^{-1} + R_y^{-1}Q_y X_y Q_y^T R_y^{-1})\right]\end{aligned} \tag{20}$$

The facts that the matrices M_x, R_x, M_y, and R_y are even functions of each one of the variables F_i ($i = 1, \ldots N$), and that the traces of the matrix expressions involving Q_x and Q_y attain their minimum value for $Q_x = Q_y = 0$ suggest that a minimum may exist for a distribution of GCP's which is symmetric with respect to the l axis.

From now on, the given number of points will be considered to be even, and the optimal distribution of GCP's will be assumed to be symmetric with respect to the l axis. (The restriction to an even number of points is made to avoid the difficulties arising from the fact that no minimum exists with $Q_x = Q_y = 0$ if some $F_i = 0$, as it will be shown later).

For ϵ_{lF} to be a minimum, a set of necessary conditions to be satisfied by the GCP's coordinates is the following

$$\frac{\partial \epsilon_{lF}}{\partial l_i} \begin{cases} \leq 0 & \text{if} \quad l_i = 1 \\ = 0 & \text{if} \quad |l_i| < 1 \\ \geq 0 & \text{if} \quad l_i = -1 \end{cases} \qquad \frac{\partial \epsilon_{lF}}{\partial F_i} \begin{cases} \leq 0 & \text{if} \quad F_i = F_M \\ = 0 & \text{if} \quad |F_i| < F_M \\ \geq 0 & \text{if} \quad F_i = -F_M \end{cases} \tag{21}$$

for $i = 1, \ldots, N$.

The derivatives of the error with respect to F_i can be written as (see Appendix)

$$\begin{aligned}\frac{\partial \epsilon_{lF}}{\partial F_i} = {} & - F_i\Big\{\sigma_x^2 \ \frac{F_M^2}{3} \ v^T(l_i)R_x^{-2}v(l_i) \\ & + \sigma_y^2 \left[2(1 + F_i^2) \left(1 + \frac{2}{3} F_M^2 + \frac{1}{5} F_M^4\right)v^T(l_i)M_y^{-2}v(l_i) + \frac{F_M^2}{3} \ u^T(l_i)R_y^{-2} u(l_i)\right]\Big\}\end{aligned} \tag{22}$$

Because the matrices R_x^{-2}, M_y^{-2} and R_y^{-2} are positive definite, Equation 22 can be put in the form

$$\frac{\partial \epsilon_{lF}}{\partial F_i} = -F_i \, H^2(l_1 \ldots l_N, F_1 \ldots F_N) \tag{23}$$

where $H^2(l_1 \ldots l_N, F_1 \ldots F_N) > 0$.

Consideration of the necessary conditions for a minimum, given in Equation 22 leads to the only three possible values of F_i: $F_i = F_M$, zero, and $-F_M$. The value $F_i = 0$ is ruled out because $(\partial^2 \epsilon_{lF}/\partial F_i^2) \, F_i = 0$ is negative, as can be seen by taking the derivative of Equation 23 with respect to F_i and setting $F_i = 0$. (Variations in F_i would lead to a decrease in ϵ_{lF}).

Therefore, the only possible optimal values left are $F_i = F_M, -F_M$; that is, the right and left edges of the image, and it should be mentioned that this result is obtained independently of the degree of the polynomials assumed for the attitude and altitude of the satellite.

Making $F_i^2 = F_M^2$, the expression for the mean square error simplifies to

$$\epsilon_{lF} = \frac{1}{2} \left[\frac{4}{3} \ \sigma_x^2 + \frac{1 + \frac{2}{3} F_M^2 + \frac{1}{5} F_M^4}{(1 + F_M^2)^2} \ \sigma_y^2\right] \text{tr}(M_x^{-1}) + \frac{\sigma_y^2}{6} \ \text{tr} \ (S_y^{-1}) \tag{24}$$

where

$$\mathbf{S}_y = \sum_{i=1}^{N} \boldsymbol{u}(l_i)\boldsymbol{u}^T(l_i)$$

The derivative of ϵ_{lF} with respect to l_i can then be written as

$$\frac{\partial \epsilon_{lF}}{\partial l} = -[\frac{4}{3}\sigma_x^2 + \frac{1 + \frac{2}{3}F_M^2 + \frac{1}{5}F_M^4}{(1 + F_M^2)^2}\sigma_y^2]\, \boldsymbol{v}^T(l_i)\, \boldsymbol{M}_x^{-2}\, \frac{d\boldsymbol{v}(l_i)}{dl_i} - \frac{\sigma_y^2}{3}\boldsymbol{u}^T(l_i)\, \boldsymbol{S}_y^{-2}\, \frac{d\boldsymbol{u}(l_i)}{dl_i} \tag{25}$$

The matrices $\boldsymbol{M}_x$ and $\boldsymbol{S}_y$ can be partitioned into submatrices in an analogous way to the one followed with $\boldsymbol{B}_x$ and $\boldsymbol{B}_y$, and the error expressed in terms of these submatrices in an entirely similar manner to the one of Equation 20, suggesting in this way that the optimal solution is also symmetric with respect to the F axis (see Appendix, paragraph A.4).

Assuming the optimal solution to be symmetric with respect to the F axis simplifies Equation 24 to

$$\epsilon_{lF} = \frac{1}{4}[\frac{4}{3}\sigma_x^2 + \frac{1 + \frac{2}{3}F_M^2 + \frac{1}{5}F_M^4}{(1 + F_M^2)^2}\sigma_y^2]\,\mathrm{tr}\,(\boldsymbol{M}_1^{-1} + \boldsymbol{M}_2^{-1}) + \frac{\sigma_y^2}{12}\left[\frac{1}{\sum_{j=1}^{N/2} P_0^2(l_j)} + \frac{1}{\sum_{j=1}^{N/2} P_2^2(l_j)}\right] \tag{26}$$

where the matrices $\boldsymbol{M}_1$ and $\boldsymbol{M}_2$ are given by

$$\boldsymbol{M}_1 = \begin{pmatrix} \sum P_0^2(l_j) & \sum P_0(l_j)P_2(l_i) \\ \sum P_0(l_j)P_2(l_i) & \sum P_2^2(l_j) \end{pmatrix}$$

$$\boldsymbol{M}_2 = \begin{pmatrix} \sum P_1^2(l_j) & \sum P_1(l_j)P_3(l_j) \\ \sum P_1(l_j)P_3(l_j) & \sum P_3^2(l_j) \end{pmatrix}$$

with all the summations taken from 1 to $N/2$, and because of the assumed symmetry $l_j = -l_{j+1}$ (if $N/2$ is odd, $l_{N/2} = 0$).

Now the problem is reduced to find the values of l_j (j odd) for which $\partial\epsilon_{lF}/\partial l_j = 0$ if $l_j < 1$ or $\partial\epsilon_{lF}/\partial l_j \leq 0$ if $l_j = 1$.

Equation 25 shows that $\partial\epsilon_{lF}/\partial l_i$ may be viewed as a polynomial of degree five in l_i, and therefore for each one of the possible local minima, at most five different values will exist for l_j making $\partial\epsilon_{lF}/\partial l_i = 0$ (by symmetry one of them is $l_j = 0$).

Numerical calculations to find the minima of ϵ_{lF} for different values of the number of points, tend to indicate that only three of the five possible different values exist; thus, the end points $l_j = \pm 1$ should be considered.

To find the absolute minimum for a given number of points the following procedure has been followed: the points are distributed among the five possible locations $l = 1, -1, 0$ and λ and $-\lambda$ keeping the assumed symmetry; λ is obtained from the equation $\partial\epsilon_{lF}/\partial\lambda = 0$, and the corresponding value for ϵ_{lF} is computed; the process is repeated for each different partition of the points into the five possible locations, and the absolute minimum is then taken from all the computed local minima. (For every local minimum it was checked that $\partial\epsilon_{lF}/\partial l_i \,|\, l_i = 1$ was negative).

In this way the minimum was found for $N = 8, 24, 32, 40$, and 48 taking $F_M = \tan 5.78°$ and $\sigma_x/\sigma_y = 79/57$ (identical standard deviations if they were expressed in pixels). The results are shown in Table 1, where it can be seen the small variation of the location given by λ; for the different values of N, the last column gives an idea of the importance of unequally distributing the points among the optimal locations.

Under a practical point of view, the solutions found suggest the areas in which GCP's should be located to lower the mean square registration error: the four corners of the image, and some areas at the right and left edges whose exact position depends slightly on the number of GCP's. The experiments described in the following paragraph show the value of this statement.

TABLE 1. OPTIMIZATION RESULTS FOR DIFFERENT VALUES OF N. THE GCP'S DISTRIBUTION IS SYMMETRIC WITH RESPECT TO BOTH AXES l AND F, AND THE F-COORDINATE TAKES ONLY THE VALUES $\pm F_M$. THE LAST COLUMN GIVES THE VALUE OF ϵ_{lF}/σ_x^2 OBTAINED BY EQUALLY DISTRIBUTING THE GCP'S AMONG THE EIGHT LOCATIONS CORRESPONDING TO THE CASE $N = 8$.

N	No. of points in each location $l = 1$	$l = \lambda$	$l = 0$	λ	$(\epsilon_{lF}/\sigma_x^2)_{\min}$	$8/N(\epsilon_{lF}/\sigma_x^2)N = 8$
8	2	2	0	0.431	0.824	0.824
24	4	8	0	0.438	0.2431	0.2747
32	6	10	0	0.436	0.1820	0.2060
40	6	14	0	0.440	0.1456	0.1648
48	8	16	0	0.438	0.1215	0.1373

REGISTRATION EXPERIMENTS

To assess the practical value of the results obtained in the preceding paragraph the following experiment was carried out:

(a) GCP's were located in a Landsat image and on the corresponding maps. The aim was to have them roughly uniformly distributed throughout the image, and after the rejection of misplaced points, 53 GCP's were kept. The mapping functions to transform the Landsat image to the UTM projection were found and approximated by two fifth-degree polynomials.

(b) Using the mapping polynomials, a set of points described by their UTM coordinates was transformed to line and pixel of the raw image with the result rounded to its nearest integer, thus producing a pseudo-location error with zero mean and standard deviation of 0.29 pixels in both directions (approximately $\sigma_x = 23.10$ m, $\sigma_y = 16.62$ m).

This simulated location of GCP's was carried out to make sure that all GCP's had the same variance because our procedure to find GCP's in an image involves the location of only one pixel in a feature (dam, airport, etc.), and this way of proceedings should make the error distribution of the GCP's dependent upon the kind of feature under consideration.

It should also be mentioned that this simulation is thought to be similar to the situation which will arise in trying to register one image to another image if cross-correlation methods are used to locate GCP's, because then the location error should be of the same order as the one coming from the simulation used here.

(c) Several sets of points, obtained as described in (b), were used as GCP's to obtain the corresponding coefficients for the assumed attitude-altitude model. In each case, the square registration error was computed at 25 uniformly distributed points of the image, using the estimated values for σ_x and σ_y given by Equation 15, and the average of these 25 values was obtained as an estimation of the mean-square error.

The simulated GCP's used for each run and the results obtained for the root mean square registration error are shown in Table 2 and Figure 2. Optimal points are those GCP's within one of the eight small rectangles depicted in the figure.

TABLE 2. RESULTS FROM THE REGISTRATION EXPERIMENTS. RUNS 1 THROUGH 4 CORRESPOND TO DISTRIBUTING THE POINTS AMONG THE EIGHT SMALL RECTANGLES SHOWN IN FIGURE 2, IN TWO DIFFERENT WAYS: (a) CORRESPONDING TO THE OBTAINED OPTIMAL SOLUTIONS GIVEN IN TABLE 1; (b) EQUALLY DISTRIBUTED. THE LAST COLUMN HAS BEEN OBTAINED USING THE VALUES GIVEN FOR $(\epsilon_{lF}/\sigma_x^2)_{\min}$ IN TABLE 1 AND $\sigma_x = 23.1$ m (THEORETICAL VALUE).

Run	No. of GCP's Opt	Nonopt	Total	$\hat{\sigma}_x$(m)	$\hat{\sigma}_y$(m)	$\epsilon_{lF}{}^{1/2}$(m)	$(\epsilon_{lF}^{1/2})_{\text{Min}}$(m)
1	24	0	24	26	15	12.09	11.39
2(a)	32	0	32	24	17	10.47	9.85
2(b)	32	0	32	23	15	10.46	
3(a)	40	0	40	23	17	9.24	8.81
3(b)	40	0	40	24	16	9.59	
4(a)	48	0	48	26	16	9.06	8.05
4(b)	48	0	48	24	16	8.64	
5	0	24	24	26	12	19.13	
6	8	16	24	26	16	14.60	
7	0	32	32	23	17	14.44	
8	24	24	48	23	18	9.45	

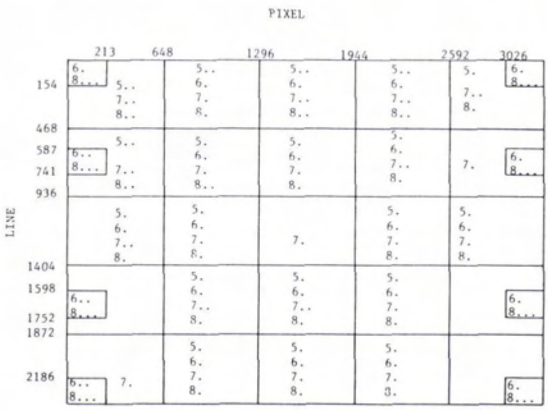

FIG. 2. Distribution of GCP's for the different runs performed. The numbers refer to the runs, and the dots to the GCP's in each rectangle.

The results of the runs compare well with the theoretical minima, considering the estimation uncertainty of the GCP's location error variance and considering also the rough calculation of the average square error.

Consideration of runs 5, 6, and 7 should give an idea of the importance of finding GCP's near the optimal locations.

CONCLUSIONS

The optimal distribution of GCP's to correct an MSS Landsat image has been found, using as an optimality criterion the mean square registration error for the whole image, and the practical value of the solution has been assessed through some experiments. The analysis, and thus the results, are not restricted to Landsat images, but are relevant to MSS images taken from a platform for which the assumptions stated in the course of the analysis remain valid. Two of these assumptions merit some further comments.

• In the introduction it was stated that all items needed to compute the geometric transformation functions were perfectly known, except for the attitude and altitude of the satellite. For Landsat images, there are important uncertainties in the given position of the nadir point, in the satellite velocity, in its orbital path, etc., that if not accounted for will cause fictitious changes in attitude and altitude in the assumed model. Furthermore, terrain relief is seldom considered in the correction of Landsat images, and its effect on the registration error could be rather important at the right and left edges of the image. All these errors are taken into account through the GCP's, and even if the estimation of the altitude and attitude of the satellite were to be degraded, the correction of the image would not be equally worsened, although obtaining the mapping functions through an attitude-altitude model would tend to lose its meaning.

In this context, it seems worthwhile to mention that errors in the position of the nadir point of the order of those encountered in practice (about 2 km) have the effect of changing the pitch and roll through the independent terms Φ_0 and ω_0 of these functions, (cf. Equations 1 and 2) because note that a bias in x can be absorbed by Φ_0 and a bias in y can be approximately absorbed by ω_0, the maximum error involved being on the order of $(\tan^2 G_{MAX}) \times$ (nadir error), $(\sim 10^{-2} \times 2 \times 10^3 \text{ m} = 20 \text{ m})$ acceptable for a practical situation.

• The second point to be commented upon is the approximation of the three attitude angles and the altitude by three cubic polynomials of time and a linear one, respectively. It seems to be a fact supported by experience (see Bernstein (1973)) that for one Landsat frame the stated approximations are adequate, meaning by that the following: the mean square error obtained by fitting a polynomial to a function assumed to be perfectly known is much smaller than the mean square error caused by the uncertainty of the values of the function. In the analysis presented here only the second type of error has been considered.

This comment is quite relevant if two or more consecutive frames of the same Landsat pass are intended to be corrected as a single image, because then the approximation of the attitude angles by cubic polynomials and of the altitude by a linear one will probably cease to be an adequate one.

In conclusion, this paper suggests that the use of models for some unknowns (necessary to compute the geometric correction functions) used in conjunction with the optimal location of GCP's can reduce the number required to obtain a given average registration error, with corresponding savings in time (human or machine).

References

Bernstein, R. 1973. Results of Precision Processing (Scene Correction) of ERTS-1 Images Using Digital Image Processing Techniques. In *Proceedings of the Symposium on Significant Results obtained from the Earth Resources Technology Satellite-1*. NASA SP-327, *Vol.* 1, Section B, pp. 1143-1150.

———, 1976. Digital Image Processing of Earth Observation Sensor Data, *IBM J. Res. Develop., Vol* 20, No. 1, pp. 40-57.

Courant, R., and D. Hilbert, 1953. *Methods of Mathematical Physics*, Vol. I. Interscience Publishers, Inc., New York.

Hamilton, W. C., 1964. *Statistics in Physical Science-Estimation, Hypothesis Testing and Least Squares*, The Ronald Press Co., New York.

Appendix

A.1. Partitions of the matrices $(\boldsymbol{W}_x^T \cdot \boldsymbol{W}_x)$ and $(\boldsymbol{W}_y^T \cdot \boldsymbol{W}_y)$.

These matrices may be written as

$$(\boldsymbol{W}_x^T \cdot \boldsymbol{W}_x) = \begin{pmatrix} \boldsymbol{M}_x & \boldsymbol{Q}_x^T \\ \boldsymbol{Q}_x & \boldsymbol{R}_x \end{pmatrix} \qquad (\boldsymbol{W}_y^T \cdot \boldsymbol{W}_y) = \begin{pmatrix} \boldsymbol{M}_y & \boldsymbol{Q}_y^T \\ \boldsymbol{Q}_y & \boldsymbol{R}_y \end{pmatrix} \tag{A.1}$$

where the submatrices have the following expressions:

$$\begin{aligned} \boldsymbol{M}_x &= \sum \boldsymbol{v}(l_i)\boldsymbol{v}^T(l_i) & \boldsymbol{M}_y &= \sum (1 + F_i^2)^2 \boldsymbol{v}(l_i)\boldsymbol{v}^T(l_i) \\ \boldsymbol{Q}_x &= \sum F_i \boldsymbol{v}(l_i)\boldsymbol{v}^T(l_i) & \boldsymbol{Q}_y &= \sum F_i(1 + F_i^2)\boldsymbol{u}(l_i)\boldsymbol{v}^T(l_i) \\ \boldsymbol{R}_x &= \sum F_i^2 \boldsymbol{v}(l_i)\boldsymbol{v}^T(l_i) & \boldsymbol{R}_y &= \sum F_i^2 \boldsymbol{u}(l_i)\boldsymbol{u}^T(l_i) \end{aligned} \tag{A.2}$$

all the summation being taken from $i = 1$ to N.

If the GCP distribution is symmetric with respect to the l axis, that is, $F_j = -F_{j+1}$, $l_j = l_{j+1}$ (j odd), then $\boldsymbol{Q}_x = \boldsymbol{Q}_y = 0$.

If all GCP's are at the right or left edges of the image, that is, $F_i^2 = F_M^2$ $(i = 1 \ldots N)$, then

$$\begin{aligned} \boldsymbol{R}_x &= F_M^2 \cdot \boldsymbol{M}_x \\ \boldsymbol{M}_y &= (1 + F_M^2)^2 \cdot \boldsymbol{M}_x \\ \boldsymbol{R}_y &= F_M^2 \sum \boldsymbol{u}(l_i)\boldsymbol{u}^T(l_i) = F_M^2 \, \mathbf{S}_y \end{aligned} \tag{A.3}$$

A.2. Inverse of a partitioned matrix. Definition of the matrices $\mathbf{X}_x, \mathbf{Y}_x, \mathbf{Z}_x, \mathbf{X}_y, \mathbf{Y}_y$, and $\mathbf{Z}_y$.

The inverse of a matrix

$$(\mathbf{W}^T \cdot \mathbf{W}) = \begin{pmatrix} \boldsymbol{M} & \boldsymbol{Q}^T \\ \boldsymbol{Q} & \boldsymbol{R} \end{pmatrix} \tag{A.4}$$

is given by

$$(\mathbf{W}^T \cdot \mathbf{W})^{-1} = \begin{pmatrix} \mathbf{X} & \mathbf{Z}^T \\ \mathbf{Z} & \mathbf{Y} \end{pmatrix} \tag{A.5}$$

where

$$\begin{aligned} \boldsymbol{X} &= (\boldsymbol{M} - \boldsymbol{Q}^T\boldsymbol{R}^{-1}\boldsymbol{Q})^{-1} = \boldsymbol{M}^{-1} + \boldsymbol{M}^{-1}\boldsymbol{Q}^T\boldsymbol{Y}\boldsymbol{Q}\boldsymbol{M}^{-1} \\ \boldsymbol{Y} &= (\boldsymbol{R} - \boldsymbol{Q}\boldsymbol{M}^{-1}\boldsymbol{Q}^T)^{-1} = \boldsymbol{R}^{-1} + \boldsymbol{R}^{-1}\boldsymbol{Q}\boldsymbol{X}\boldsymbol{Q}^T\boldsymbol{R}^{-1} \\ \boldsymbol{Z} &= -\boldsymbol{R}^{-1}\boldsymbol{Q}\boldsymbol{X} = -\boldsymbol{M}^{-1}\boldsymbol{Q}^T\boldsymbol{Y} \end{aligned} \tag{A.6}$$

The matrices $\boldsymbol{X}_x, \boldsymbol{Y}_x, \boldsymbol{Z}_x, \boldsymbol{X}_y, \boldsymbol{Y}_y$, and $\mathbf{Z}_y$ introduced in Equation 18 are given by Equation A.6 with the corresponding subindeces x or y.

A.3. Derivatives of the traces of the matrices $\boldsymbol{M}_x^{-1}, \boldsymbol{R}_x^{-1}, \boldsymbol{M}_y^{-1}, \boldsymbol{R}_y^{-1}$.

Using the derivation rules concerning the inverse of a matrix

$$\frac{\partial}{\partial F_i} \operatorname{tr}(\boldsymbol{R}_x^{-1}) = -\operatorname{tr}\left(\boldsymbol{R}_x^{-1} \frac{\partial \boldsymbol{R}_x}{\partial F_i} \boldsymbol{R}_x^{-1}\right) \tag{A.7}$$

Taking the derivative of $\boldsymbol{R}_x$ in Equation A.2

$$\frac{\partial}{\partial F_i} \quad \mathrm{tr}(\boldsymbol{R}_x^{-1}) = -\mathrm{tr}(\boldsymbol{R}_x^{-1}.2.F_i\boldsymbol{v}(l_i)\boldsymbol{v}^T(l_i)\boldsymbol{R}_x^{-1}) \tag{A.8}$$

Changing the order of the matrix product within the parenthesis

$$\frac{\partial}{\partial F_i} \mathrm{tr}(\boldsymbol{R}_x^{-1}) = -2F_i \, \boldsymbol{v}^T(l_i)\boldsymbol{R}_x^{-2} \, \boldsymbol{v}(l_i) \tag{A.9}$$

Similarly, it can be written

$$\frac{\partial}{\partial F_i} \mathrm{tr} \, (\boldsymbol{M}_y^{-1}) = -4F_i(1 + F_i^2)\boldsymbol{v}^T(l_i)\boldsymbol{M}_y^{-2} \, \boldsymbol{v}(l_i)$$

$$\frac{\partial}{\partial F_i} \mathrm{tr} \, (\boldsymbol{R}_y^{-1}) = -2F_i \, \boldsymbol{u}^T(l_i)\boldsymbol{R}_y^{-2}\boldsymbol{u}(l_i) \tag{A.10}$$

In a completely analogous way, the derivatives of the traces of $\boldsymbol{M}_x^{-1}$ and $\boldsymbol{S}_y^{-1}$ with respect to l_i can be obtained

$$\frac{\partial}{\partial l_i} \mathrm{tr} \, (\boldsymbol{M}_x^{-1}) = -2 \, \boldsymbol{v}^T(l_i) \, \boldsymbol{M}_x^{-2} \, \frac{\mathrm{d}\boldsymbol{v}(l_i)}{\mathrm{d}l_i} \tag{A.11}$$

$$\frac{\partial}{\partial l_i} \mathrm{tr} \, (\boldsymbol{S}_y^{-1}) = -2 \, \boldsymbol{u}^T(l_i) \, \boldsymbol{S}_y^{-2} \, \frac{\mathrm{d}\boldsymbol{u}(l_i)}{\mathrm{d}l_i}$$

A.4. Partition of the matrices $\boldsymbol{M}_x$ and $\boldsymbol{S}_y$.
Let

$$\boldsymbol{v}_{EV}^T(l_i) = [P_0, (l_i), P_2(l_i)],$$
$$\boldsymbol{v}_{OD}^T(l_i) = [P_1(l_i), P_3(l_i)],$$

thus

$$\boldsymbol{v}^T(l_i) = [\boldsymbol{v}_{EV}^T(l_i), \boldsymbol{v}_{OD}^T(l_i)]$$

then

$$\boldsymbol{M}_x = \sum \boldsymbol{v}(l_i)\boldsymbol{v}^T(l_i) = \begin{pmatrix} \sum \boldsymbol{v}_{EV}(l_i)\boldsymbol{v}_{EV}^T(l_i) & \sum \boldsymbol{v}_{EV}(l_i)\boldsymbol{v}_{OD}^T(l_i) \\ \sum \boldsymbol{v}_{OD}(l_i)\boldsymbol{v}_{EV}^T(l_i) & \sum \boldsymbol{v}_{OD}(l_i)\boldsymbol{v}_{OD}(l_i) \end{pmatrix}$$
$$= \begin{pmatrix} \boldsymbol{m} & \boldsymbol{q}^T \\ \boldsymbol{q} & \boldsymbol{r} \end{pmatrix}$$

and

$$\mathrm{tr} \, (\boldsymbol{M}_x^{-1}) = \mathrm{tr} \, (\boldsymbol{m}^{-1} + \boldsymbol{r}^{-1} + \boldsymbol{m}^{-1} \, \boldsymbol{q}^T(\boldsymbol{r} - \boldsymbol{q}\boldsymbol{m}^{-1}\boldsymbol{q}^T)^{-1}\boldsymbol{q}\boldsymbol{m}^{-1} + \boldsymbol{r}^{-1}\boldsymbol{q} \, (\boldsymbol{m}^{-1} - \boldsymbol{q}^T\boldsymbol{r}^{-1}\boldsymbol{q})^{-1}\boldsymbol{q}^T\boldsymbol{r}^{-1})$$

which shows the similarity with Equation 20, as the matrices $\boldsymbol{m}$ and $\boldsymbol{r}$ are even functions with respect to each one of the l_i, because $\boldsymbol{v}_{EV}(l_i)$ is an even function and $\boldsymbol{v}_{OD}(l_i)$ is odd. The matrix $\boldsymbol{S}_y$ is treated analogously.

(Received 29 September 1978; revised and accepted 26 July 1980)

COVER PHOTOS NEEDED