BRUCE FORSTER
Department of Photogrammetry
University of New South Wales
Sydney, Australia

Urban Control for Landsat Data

A stepwise polynomial transformation of 100 ground control points over the Sydney Metropolitan area produced transformation parameters which achieved standard errors of approximately 30 metres.

ABSTRACT: *The importance of accurate registration of ground truth areas to equivalent landsat digital data in urban areas must be stressed, if detailed classification is to be achieved. A stepwise polynomial transformation of 100 ground control points over the Sydney Metropolitan area produced transformation parameters which achieved standard errors of approximately 30 m. The least-squares polynomial transformation was computed using a standard multiple regression statistical package. Shade prints and character prints of typical ground control points were generated to illustrate a small feature selection method. The effect of the ultimate transformation on the control points was examined visually on a plot of the residuals.*

INTRODUCTION

CURRENTLY throughout the world, many studies are using remotely sensed data from the Landsat satellite to classify urban areas. One of the results of this work is a proposed standard classification system (Anderson and Witmer, 1976) for urban areas suitable for use with remote sensing. Level I is Urban (or built up areas) and Level II is a further breakdown into Residential, Commercial, Industrial, etc. Only limited success has been achieved classifying Landsat data into Level II.

Unlike cultivated rural areas, urban areas exhibit an extremely heterogeneous surface cover, and considerable inter-pixel and intra-pixel change can occur. Ground truth areas therefore need to be spatially related to the Landsat digital image at the sub-pixel level if the correlation between ground and image is to be fruitfully examined.

Obviously the number, distribution, and registration accuracy of ground control points (GCP) will influence the accuracy of the computed ground image spatial relationship. Bernstein (1976) shows that if GCP registration accuracy is 0.5 of a pixel then 12 GCP are needed to obtain a 50 metre root-mean-square error and 18 GCP for a 40 metre root-mean-square error. As Carter (1977) notes, these theoretical levels are not so easy to obtain.

This paper reports the results of using 100 GCP distributed over the Sydney, Australia metropolitan area. This area forms part of Landsat scene No. 1141-23140, and covers a square of approximately 35 km sides, or approximately 600 Landsat pixels east/west by 450 pixels north/south.

TRANSFORMATION PROCEDURES

Current research on the selection of suitable formulae for the geometric correction of Landsat data reveals three basic methods (Trinder 1978):

- A parametric solution in which an attempt is made to determine the position and attitude of the spacecraft using known for-

PHOTOGRAMMETRIC ENGINEERING AND REMOTE SENSING,
Vol. 46, No. 4, April 1980, pp. 539-545.

mulae describing the image formation process;

- Polynomial interpolation formulae which are a simplified approach to the adjustments, and do not require any knowledge of the image formation process; and
- A simple interpolation formula, followed by a least-squares linear interpolation procedure.

Trinder (1978) suggests that there appears to be little if any difference between results obtained using all three methods above. Because of its simplicity and ease of application using an available multiple regression computer package, the polynomial method was chosen for this investigation.

The adjustment equations initially chosen were complete 5th order polynomials of the form

$$E = a_0 + a_1x + \ldots + a_5x^5 + a_6y + \ldots + a_{10}y^5 + a_{11}xy + \ldots + a_{23}x^4y^3 \ldots \quad (1)$$

$$N = b_0 + b_1x + \ldots + b_5x^5 + b_6y + \ldots + b_{10}y^5 + b_{11}xy + \ldots + b_{23}x^4y^3 \ldots \quad (2)$$

where E and N are East and North coordinates respectively.

(These coordinates were Integrated Survey Grid (ISG) coordinates, a local system based on a transverse mercator projection, which is used for large scale mapping in the state of New South Wales);

x and y are the pixel and line numbers respectively of the control points; and

a_0 to a_{23} and b_0 to b_{23} are the transformation parameters.

In addition to these transformations, a correction for interswath discontinuity was applied. This cannot be corrected by the polynomial transformations. The discontinuity, or apparent slip between swathes, is due to both Earth rotation and sensor delay effects, and is on the order of 50 m for Australian latitudes (Trinder, 1978). A 50-m shift is applied to all lines in each consecutive swath of six lines; the remaining linear skew is then accounted for by the polynomial transformation.

As the polynomial equations are essentially a least-squares adjustment with E and N as the dependent variables and x and y (and higher orders) as the independent variables, the problem can be treated using multiple regression techniques to compute the transformation parameters.

Such a program was a standard package available on the University of New South Wales CDC Cyber computer and is one of a set of programs in the Statistical Package for the Social Sciences (SPSS) (Nie *et al*, 1975).

Each polynomial term is computed and treated as a linear variable for the purpose of the regression. This particular program allowed for stepwise regression. Here the independent variables are entered one by one, the order of inclusion being determined by the respective contribution of each variable to explained variance. In addition, variables are deleted if they no longer meet the next best criterion. At any given step, therefore, the best sub-set of variables is used in the regression equation.

For each coefficient (or transformation parameter) the numerical value, standard error, F test, and level of significance are printed. For the regression equation as a whole, R square (or percentage explanation), F test, level of significance, and standard deviation are the major statistical data printed. For each control point, actual coordinate values, estimated values, and residuals can also be tabulated.

From 23 transformation parameters, the stepwise method allows the selection of the optimum variable set, after which little improvement in terms of standard error is gained. The optimum set can amount to as few as three or four variables.

In addition to the interswath geometric error corrected prior to the regression procedure, other significant geometric errors may exist. These include the effects of non-linear scan rate of the oscillating scanning mirror and a topographic effect due to height differences of the various GCP. A full description of all geometric errors has been documented elsewhere (Kratky, 1975).

For the selected Landsat sub-scene, topographic effect is insignificant since the maximum height difference is only 200 m, and maximum distance from the satellite nadir is approximately 500 pixel widths. Using $\Delta r = \Delta z \; r/z$ (where z is the elevation of the satellite and r is the pixel coordinate along a particular scan line as measured from the satellite nadir), the maximum possible correction is only on the order of 0.1 pixel widths or 6 m.

The effect of non-linear scan rate can generally be accounted for by higher order polynomials. In addition, the selected sub-scene extends east only 30 km from the satellite nadir and, thus, the main components of this effect will be approximately linear. Trinder (1978) found that the non-linear scan rate effect was completely eliminated by the 3rd order terms in a complete 3rd order polynomial.

Control Point Selection and Distribution

The purpose of the investigation was to allow transformation between parts of the Landsat Sydney scene and the ground truth areas. These ground truth areas were located at approximately 4 km centers, and sufficient ground control to ensure a tight fit to each area was required.

Computer shadeprints and a corresponding printout of the intensity values could be produced for any specific area. To produce the shadeprints, a histogram of the intensity values was determined and nine different gray scale values were assigned to equal portions (Trinder, 1978).

The corresponding 'map' of the intensity values was achieved using a character to represent all occurrences of a decimal value. For efficient computation and storage on the CDC Cyber computer, the original digital data in a seven bit code was reduced to a six bit code, with an effective halfing of the range of response values in each band. For Band 7 the response range (0 to 32) is represented by the characters A to Z for the values 1 to 26, 0 to 5 for the values 27 to 32, and semicolon (;) represents a zero response.

An examination of various shadeprints from each of the four bands indicated that, for urban areas, Band 7 was preferable for GCP selection. This is due to the abundance of small parks, intersecting man-made linear features, small water bodies, and, for coastal cities, natural and man-made projections into a water body. All of these features are generally against backgrounds which give significant contrast in this near infrared band.

Control points were selected from within areas of 50 by 50 pixel dimension. While this produced a more dense distribution of control points in the along scan direction, since Landsat pixels are 1.4 times longer in the across scan direction, this was considered an advantage as most non-linear scanning errors tend to have their maximum effect in this direction. The selection of the best GCP in each area also meant that in some cases the points were relatively close together.

Initial control point selection was made using the shade prints, while final estimation of pixel coordinates was achieved using the character printout. Pixel and line values for each GCP were estimated to 0.5 of a pixel. For small ground features on uniform

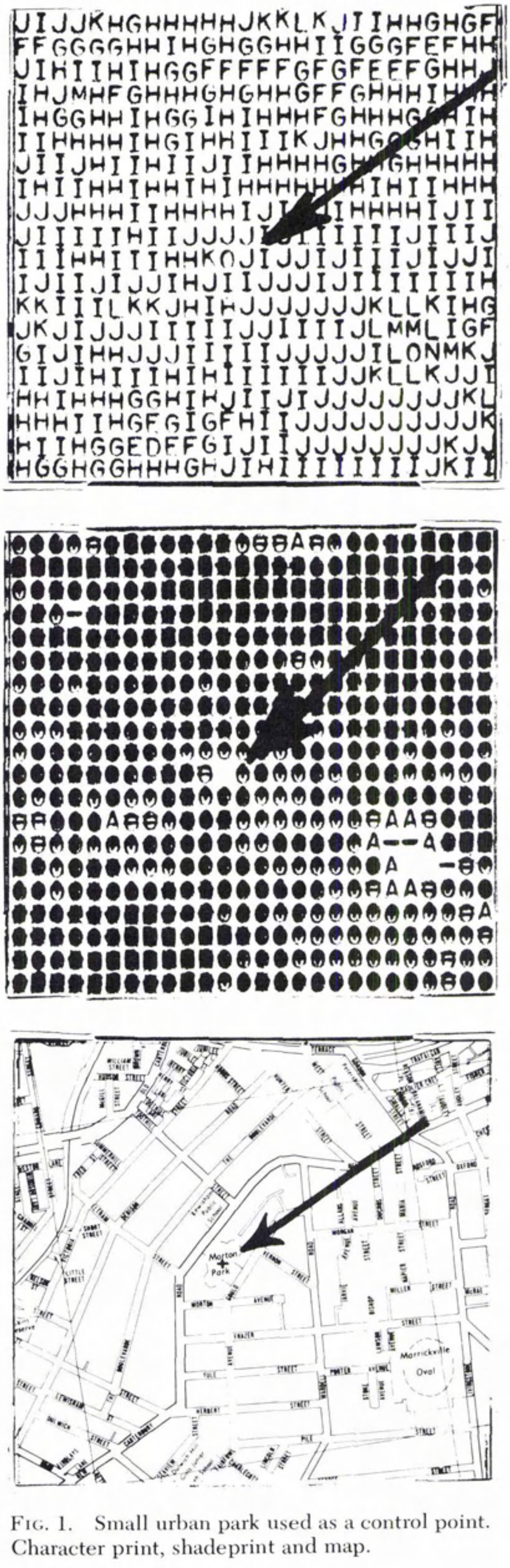

Fig. 1. Small urban park used as a control point. Character print, shadeprint and map.

backgrounds, the displays reveal a symmetric distribution of intensities if the pixel coincides exactly with the detail. Interpolation in these cases is relatively easy, giving whole or half pixel values. If the ground feature is asymetric with respect to the pixel location, the distribution of intensity values is also asymetric and an estimation of the location to the nearest half pixel is required. This can lead to a maximum error of ± 0.25 pixel.

Pixel and line values have their origin in the north-west corner of the Landsat frame, positively increasing west to east for pixel values and north to south for line values. The Sydney sub-scene begins at approximately line 1500 and pixel 1300.

The coordinates of the GCP were obtained from 1:10,000 series planimetric maps which were available for the whole of the Sydney region. Coordinates were estimated to the nearest 5 m by referring to the closest grid intersection and measuring the small displacements from that grid intersection.

Preferred control points were features on the order of one or two pixels in size, against a contrasting background. The center or intersection of features was chosen in preference to the edge where the intensity change takes place. Unless enhanced, edges are smeared due to the low resolution of the Landsat scanner and an unambiguous interpretation is difficult to achieve.

Four examples of typical urban control points are shown in Figures 1 to 4. Each figure shows the character print, shadeprint, and a 1:20,000 (original 1:10,000) scale map portion of the equivalent area.

Figure 1 is a small urban park of dimensions slightly larger than one pixel. It illustrates how a small feature with a strongly contrasting background is easily located on the shadeprint. The contrasting unique pattern of the urban cover also allows for unambiguous identification of the feature. Here the larger Marrickville oval to the southeast confirms the identification. Examination of aerial photographs or photo mosaics, in addition to the map, is preferred for control point selection. Many small open spaces and other significant features visible on the shade prints are not specifically mapped. Problems of recognition and identification of control points without aerial photographs were emphasized by Carter (1977).

The intersection of two main airport runways is shown in Figure 2. Because of the width of the runway, it is difficult to identify the control point from the shadeprint. The character print is unambiguous, however,

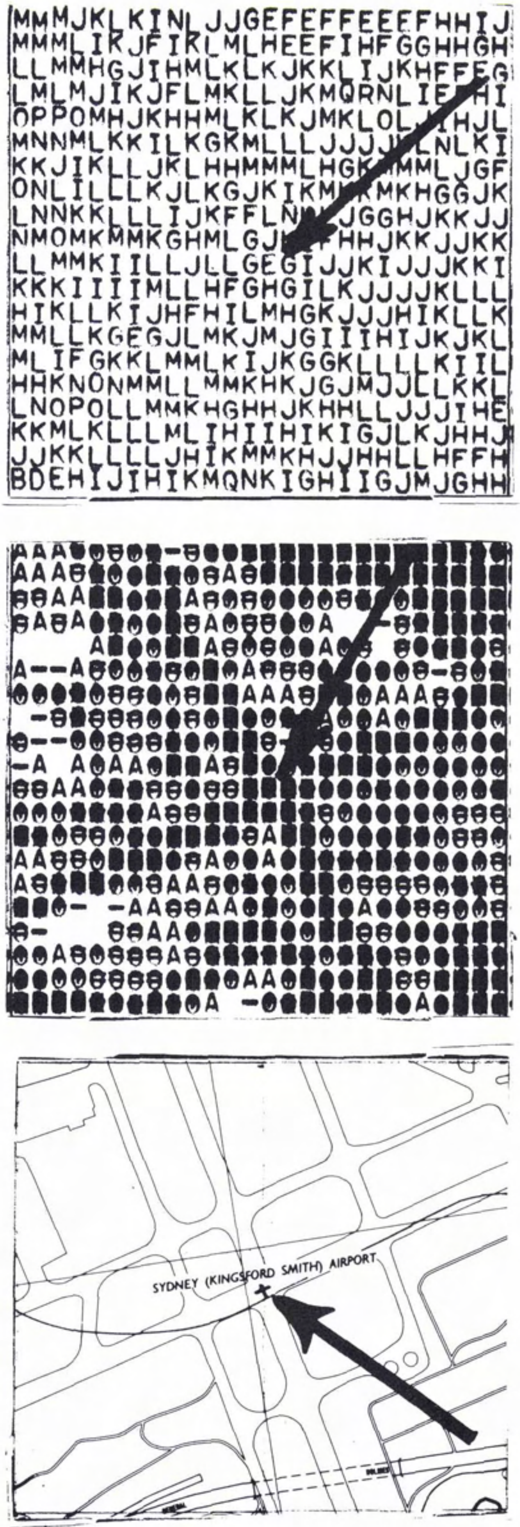

Fig. 2. Intersection of airport runways used as a control point. Character print, shadeprint, and map.

giving minimum response at the intersection, corresponding to the character E.

The capability of Landsat to record a response from narrow linear features is clearly indicated in Figure 3. The jetty projecting into the bay is only 20 m wide, yet a significant contrast is present. This shows clearly on the character print as a series of B and C values against a lower response background. The major difficulty with this type of feature is the interpretation of the intersection of the shore line with the linear feature. While in most cases a steep response gradient exists between water and land, tide conditions at the instant of scanning are important where the water-land interface is a graded beach. Preferred intersections of natural or man-made features with coastline are those where the intersection is positionally constant, as with cliffs.

Figure 4 illustrates the reverse of Figure 3. Here a small water feature is surrounded by a land mass. In this case the feature is a small lake situated in a golf course. This situation is optimum as the surrounding vegetation gives a high response in the near infrared, while the water feature gives a low to zero response. Due to the low resolution of the sensor, a local minimum considerably above zero is present. Against a dense man-made background, this size water feature would not be distinguishable.

Figure 4's lake example also points up the care one must take to avoid gross errors. Initially the lake feature to the south of the central point was marked and accepted as the corresponding ground point. Coordinates for this point were taken from the map sheet. An initial computation of the transformation parameters revealed a residual of the order of 200 m at this GCP. After checking the ground and Landsat coordinates, to no avail, the original photographs upon which the map series was based were checked. By comparing surrounding road features it was found that, on the map, the lake was plotted 200 m out of position. The lake position was corrected on the map and the true coordinates of its center were adopted.

Results of Computation

Standard errors of the order of 30 m were achieved for both Equations 1 and 2. The distribution of vector residuals is shown in Figure 5. At map scale each cell represents 2850 by 3950 metres; therefore, each cell contains 50 by 50 Landsat elements. For convenience, each control point is located at the center of the cell from which it was selected. The vector residuals at each con-

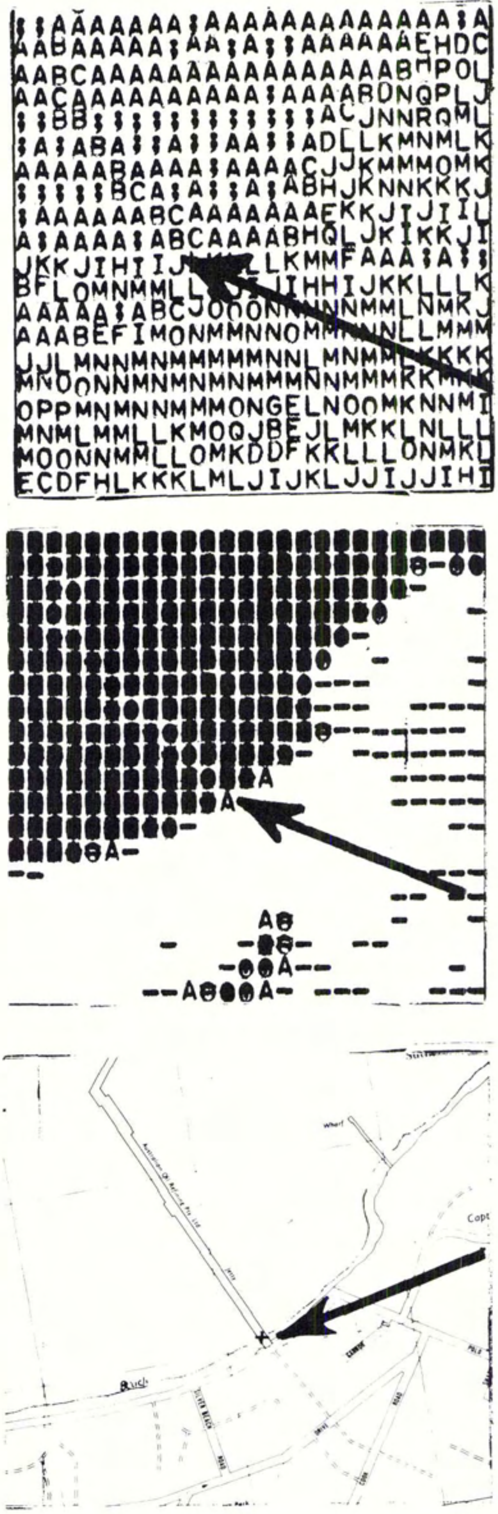

Fig. 3. Intersection of a jetty with a beach used as a control point. Character print, shadeprint and map.

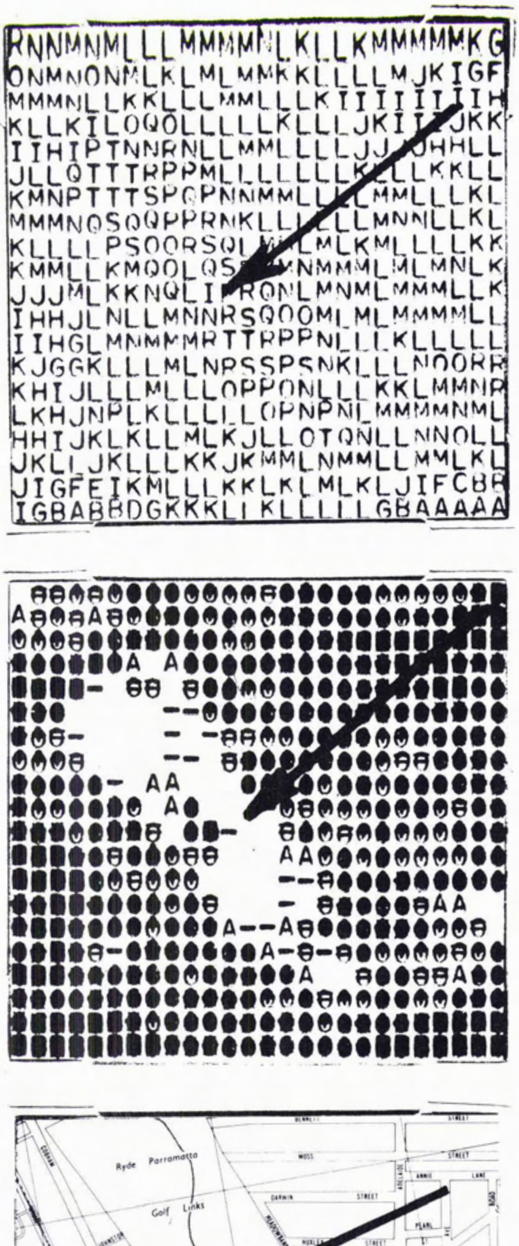

FIG. 4. Small urban lake used as a control point. Character print, shadeprint and map.

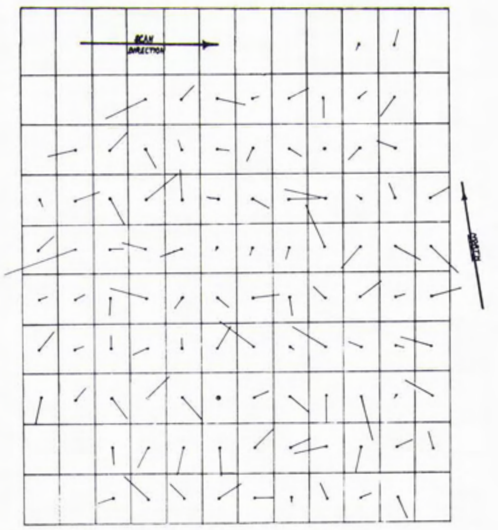

FIG. 5. Vector residual distribution 100 GCP. Each cell is 57 by 79 m at vector scale. The centers of each cell are also taken to represent the approximate relative position of the control points on the ground, being separated by 2850 m along scan and 3950 m across scan.

trol point are drawn at a 50 times enlargement, and so at this scale the cell represents a 57 by 79 metre picture element.

To reduce the size of the values used in the regression analysis, 250,000 metres and 1,200,000 metres were subtracted from the Zone 56 ISG. Easting and Northing coordinates, respectively. Line and pixel values were reduced by 1000 units and divided by 10. Thus, in the resulting equations, x and y values are fractions in units of 10 pixels or lines. By scaling down the values to be used in the computations, computer round-off error is minimized.

For Equation 1 the significant variables, at a 90 percent confidence level or greater, were x, y, and x^5, all other variables being insignificant. The standard error of the regression as a whole using these three variables was 32.3 m, and the equation was significant at the 99.9 percent level. The order of entry of the variables into the regression computation was x, y, and x^5, with the standard error improving from 1376 m to 33 m as the y value entered the equation, and only a marginal but significant improvement with the entry of x^5. The coefficient of x^5 was significant at the 98.6 per cent level.

More variables were required to define adequately the northing transformation. Here the significant variables were, in order of entry, y, x, xy^2, x^5, and x^3. The resulting standard errors, as each new variable was entered, were 1848 m, 26.34 m, 26.29 m, 26.38 m, and 24.79 m. Little if any improvement in the standard error was achieved as variables xy^2 and x^5 were entered, even though their coefficients were significant at the 90 percent level or better. Least significant however was xy^2 in the final variable set, and this could possibly be removed with little change in the results. Overall equation significance was greater than 99.9 per cent.

The final results for each equation were

$$E = 43564.7 + 569.19x - 111.69y - 0.27x^5 (10^{-7})$$

and

$$N = 126705.3 - 794.07y - 108.57x + 0.14xy^2 (10^{-3}) + 0.18x^5 (10^{-6}) - 0.15x^3 (10^{-2})$$

where E and N are in metres and x and y are in units of ten pixels or lines, respectively.

Similar equations with x and y as the dependent variables and polynomials in E and N as the independent variables can also be derived for use when transforming ISG coordinates into the image coordinate system.

Reducing the number of GCP to 50 by selecting every alternate GCP produced virtually the same results, a confidence level of better than 99.9 percent for the equations as a whole, and standard errors of the same order. This suggests that, with careful selection of GCP, the theoretical result of Bernstein (1976) can be reached.

Because the use of a fifth order polynomial can lead to extensive warping, especially at the edges of the control block, the derived equations should only be used to predict coordinates of points well within the perimeter.

Conclusion

Transformation between ground coordinates and Landsat line and pixel coordinates was achieved using a standard multiple regression statistical package. The Landsat Band 7 was found to be preferable for control point selection because of the high contrast between man-made, vegetated surfaces, and water surfaces in this band. Identification of the control point on the ground was made easier if aerial photographs of the area were available in addition to large scale base maps.

The results of the polynomial transformation—standard errors of approximately 30 m for both Northings and Eastings—suggest, for urban areas at least, that a sub-pixel spatial correspondence between ground truth areas and observed response can be achieved. This would allow a more detailed examination of the urban response and a greater confidence in classification results.

References

Anderson, J. R., and Whitmer, 1977. The use of remotely sensed data in the land use and land cover map and data compilation programme or the U.S. Geological Survey. Presented Paper, European Seminar on Regional Planning and Remote Sensing, June 1977, Toulouse.

Bernstein, R., 1976. *Digital Image Processing of Earth Observational Sensor Data,* IBM Journal of Research and Development, 20 No. 1. 40-57.

Carter, P., 1977. *Automated Land Use Mapping Using Aerial/Satellite Data,* Department of Environment Report, RRL 77/543, London.

Kratky, V., 1975. Precision Processing of ERTS Imagery. Presented Paper, ACSM-ASP Semi-Annual Convention, 1975, San Francisco.

Nie, N. H., C. Hadlai Hull, J. G. Jenkins, K. Steinbrenner, and D. H. Bent, 1975. *Statistical Package for the Social Sciences,* (Second edition) McGraw-Hill, New York. 675p.

Trinder, J. C., 1978. Rectification of Landsat Digital Data. Invited Paper, U.S.—Australia Workshop on Image Processing Techniques for Remote Sensing, May 1978, Canberra.

(Received 23 February 1979; revised and accepted 7 November 1979)

Back Issues of the Journal Available

The U.S. Geological Survey, Reston, Virginia has the complete volumes of *Photogrammetric Engineering and Remote Sensing* from 1961 through 1979 which they wish to donate to an educational institution. Please call Mrs. Peters at (703) 860-6276.