Rectification of Digital Imagery

Kurt Novak

Department of Geodetic Science and Surveying, The Ohio State University, 1958 Neil Avenue, Columbus, OH 43210

ABSTRACT: Different methods can be applied to generate digital orthophotos. Three commonly used approaches are analyzed and compared in this paper. They can be applied to rectify both digitized aerial photographs and satellite scenes. These methods are polynomial, projective, and differential rectifications. The first two are defined by analytical transformations between image and orthophoto without considering the geometry and orientation of the camera. They are approximate solutions. The last one models the physical reality of the imaging process by means of the collinearity equations and corrects for relief displacements. For eliminating distortions of the camera system, additional parameters were included. This proves to be necessary when images are directly taken by video cameras. All three methods were implemented on a workstation and were tested with digitized aerial photographs and video images. By overlaying GIS data over the digital orthophoto, the quality of the rectification is checked. To determine the planimetric accuracy of the results, the coordinates of targets were measured in a digital orthophoto and compared to known map coordinates.

INTRODUCTION

The combination of regular line maps with aerial photographs is desirable for many applications. While line maps usually consist of topographic elements such as roads, buildings, and rivers, many features important for disciplines like geology, archaeology, forestry, etc., are missing. However, they can be found in aerial or satellite images. In order to be able to exactly overlay images and maps, their geometry has to be the same: that is, an orthogonal projection of all points of the ground to a reference surface. This process is called *rectification*, and the resulting image-based map is the *orthophoto*.

Recently rectification methods have gained more significance due to the fact that digital maps (stored in a geographic information system, GIS) are to be combined with digital imagery. To make sure that they are overlaid at the correct locations, both have to be related to the same datum (map projection). Many image processing systems designed for digital mapping and remote sensing provide routines that perform this rectification procedure; it is sometimes called geo-referencing or geocoding. The mathematical models applied range from simple affine transformations, utilizing higher order polynomial and projective transformations, to differential rectification with relief displacement correction.

The last method is of particular importance if a surface model of the landscape has to be created. Because digital images and digital elevation data are both related to a planimetric grid, they are basically stored in the same way: as matrices of gray or elevation values. Digital rectification assigns a gray value to each grid-element of the digital elevation model (DEM), so that both elevation and density of the surface are stored at the same planimetric location.

This paper investigates different rectification methods, explains the modifications necessary to apply them to satellite imagery, and tries to point out the advantages and setbacks of each. Practical examples show how the planimetric positions of points vary in the different orthophotos. Differential rectification is enhanced by additional parameters in order to correct for lens and film distortions. This is important if video cameras are used for data collection, as well as with satellite imagery to model deviations of the orbit.

Only a few authors have covered digital rectification in recent years. The first system for digital orthophoto generation using scanning microdensitometers was presented by Keating and Boston (1979). This system replaces the optical projection of a small slit through zoom lenses and dove prisms (applied in analytical orthophoto projectors) by electronic sensors that digitize small parts of the image, a processor that positions the sensor, and a film writer that directly exposes each pixel on the film. The basics of digital rectification are explained by Konecny (1979), and are discussed in Baehr (1985). A very interesting method has been implemented by Jansa (1983). He uses linear prediction to recitfy Landsat MSS images in a dense network of control points. Digital rectification is an important module of digital stereo workstations (Mueller *et al.*, 1988; Hood *et al.*, 1989). Systems of that kind are mostly based on commercial image processing systems, and allow one to combine digital orthophotos with vector data of a GIS.

TRANSFORMATION OF DIGITAL IMAGES

To transform a digital image by an analytical function, one can pursue either a direct or an indirect approach. The *indirect method* takes each pixel location of the result (e.g., orthophoto), determines its position in the original image by the selected transformation f_x and f_y (Equation 1), and interpolates the gray value by a given resampling method (Figure 1a); that is,

 $x = f_x(x', y')$ indirect transformation of point (x', y') (1) $y = f_y(x', y')$ from the result to the original image.

where

x, y = pixel coordinates of the original image, x', y' = pixel coordinates of the result, and $f_x(), f_y()$ = transformation functions.

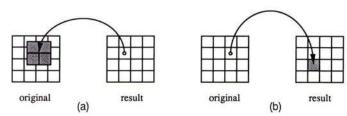


Fig. 1. Indirect (a) and direct (b) transformations of a pixel.

On the other hand, the *direct method* starts from the pixel location in the original image, transforms its coordinates into the result, and places the gray value to the nearest integer pixel

(Figure 1b).

The direct method only assigns gray values that are given by the original image; therefore, the contrast and density are not changed by the transformation. However, it might happen that some pixels do not get any gray value, which means that they have to be filled after the transformation in a second pass through the resulting image. The indirect transform assigns a gray value to exactly one pixel in the result (an integer coordinate pair). It has to be computed by interpolation from the original gray values. This is necessary, as the location of the pixel which was transformed into the original does not correspond to an integer position. The process is called *resampling*. By using the indirect method, the gray values are altered in dependency of the resampling algorithm.

The resampling method is chosen with respect to the accuracy, contrast, and resolution required in the result. The ideal resampling function is the sinc-function $(\sin(x)/x)$. It cannot be applied practically, as an infinite number of pixels would have to be convolved with the original image to compute one gray value of the result. Therefore, it is approximated by simpler resampling functions. The most popular ones are listed in Table 1, which compares the size of the convolution window and the number of multiplications necessary for the interpolation of one

pixel.

For most applications bilinear resampling yields sufficiently good results. The cubic convolution gives a better contrast than bilinear resampling, which smoothes the original gray values. Nearest neighbor resampling causes step patterns in almost horizontal or vertical lines; however, it is the least computationally intensive method.

THREE RECTIFICATION METHODS

Three methods were investigated for this paper. They are commonly used in image processing systems, and represent only part of the spectrum of approaches available.

POLYNOMIAL RECTIFICATION

The transformation between the original and the rectified images is done by polynomials (Equation 2). That is,

$$x = \mathbf{x}'^{T} \mathbf{A} \mathbf{y}' = f_{x} (x', y')$$

 $y = \mathbf{x}'^{T} \mathbf{B} \mathbf{y}' = f_{y} (x', y')$ (2)

where

x,y are coordinates of the original image, x', y' are coordinates of the rectification, and A, B are coefficient matrices of the polynomials.

$$\mathbf{x}^{\prime T} = (1, x', x'^{2}, x'^{3}, \dots)$$

$$\mathbf{y}^{\prime T} = (1, y', y'^{2}, y'^{3}, \dots)$$

$$\mathbf{A} = \begin{pmatrix} a_{00} & a_{01} & a_{02} & \dots \\ a_{10} & a_{11} & a_{12} & \dots \\ a_{20} & a_{21} & a_{22} & \dots \end{pmatrix} \quad \mathbf{B} = \begin{pmatrix} b_{00} & b_{01} & b_{02} & \dots \\ b_{10} & b_{11} & b_{12} & \dots \\ b_{20} & b_{21} & b_{22} & \dots \end{pmatrix}$$

TABLE 1. COMPARISON OF THREE DIFFERENT RESAMPLING METHODS.

Method	Window	Multiplications
Nearest Neighbor Resampling	$(1 \times 1 \text{ pixel})$	0
Bilinear Resampling	$(2 \times 2 \text{ pixels})$	4
Cubic Convolution	$(4 \times 4 \text{ pixels})$	110

This method corrects for distortions of the image relative to a dense set of control points. The order of the polynomial depends on the number of control points available. The more control information one has, the more accurate the result of the rectification will be. This approach is completely independent of the geometry of the imaging sensor, so it can be used both for satellite images and aerial photographs. Due to the polynomial transform, the original image is shifted, rotated, scaled, and squeezed so that it fits best to the given reference points. This is the reason why the method is most often used for satellite images, whose geometry and distortions are sometimes difficult to model. Moreover, the relief displacement due to the topography of the Earth is relatively small compared to the flying height of the satellite and does not influence the result significantly.

Although polynomials are very easy to use for rectifications, they can cause problems and errors in the transformed image. They do not adequately correct relief displacements, nor do they consider the special geometry of the imaging system. For the inexperienced user a high-order polynomial may seem to provide a perfect fit at the reference points, as the residuals are small; however, due to undulations between the reference points, large errors could be created. The biggest advantage of the polynomial transform is the fact that all distortions of the image (due to sensor geometry, relief displacement, Earth curvature, etc.) are corrected simultaneously.

PROJECTIVE TRANSFORMATION

The projective transform describes the relationship between two planes (Figure 2). It is defined by eight parameters (Equation 3), which can be derived from four object points lying in a plane and their corresponding image coordinates. The elements of exterior and interior orientations are not needed, as they are implicit in these parameters. This method is typically used to rectify aerial photographs of flat terrain or images of facades of buildings.

$$x = \frac{a_1 x' + a_2 y' + a_3}{c_1 x' + c_2 y' + 1} = f_x(x', y')$$

$$y = \frac{b_1 x' + b_2 y' + b_3}{c_1 x' + c_2 y' + 1} = f_x(x', y')$$
(3)

where a_1 , a_2 , a_3 , b_1 , b_2 , b_3 , c_1 , c_2 are the projective parameters.

The projective transformation can be modified so that satellite imagery can be rectified. Satellite sensors, such as SPOT, only capture one line of the image at a time. Therefore, we get a

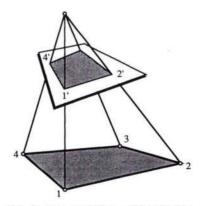


Fig. 2. The projective relationship between an object and the image is defined by eight parameters which can be determined from four control points.

line-perspective relationship between the ground (plane) and the image (Figure 3). Each scan-line has a different perspective center, and one can assume that all are tied to a straight line approximating the orbit. If y is given in the flying direction and x represents the pixel in a scan line, the projective transform is modified to Equation 4. That is,

$$x = \frac{a_1 x' + a_2 y' + a_3}{c_1 x' + c_2 y' + 1}$$

$$y = b_1 x' + b_2 y' + b_3$$
(4)

Due to Earth curvature, one can hardly define a plane area on the ground which can be related to a satellite scene by Equation 4. Even for very flat terrain the Earth curvature cannot be

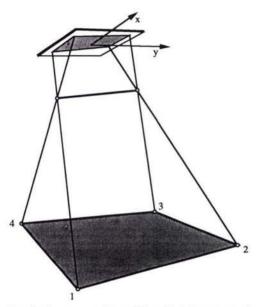


Fig. 3. Line-perspective relationship between an object plane and a satellite scene.

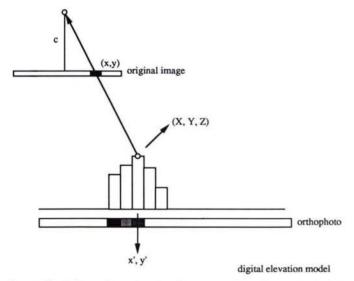


Fig. 4. Each three-dimensional surface point defined as a pixel of the DEM is transformed into the image to find the density from a digitized aerial photograph. This gray value is assigned to the orthophoto raster at the same pixel location as the DEM point.

neglected. Therefore, this method has little practical significance for satellite scenes, but could be applied for airborne line

DIFFERENTIAL RECTIFICATION

The term differential rectification has its origin in the approach of rectifying small parts as of a photograph at a time. Conventional analog or analytical orthoprojectors do this by optical image transfer. A tiny slit is moved over the orthophoto, while its corresponding part in the original image is found by shifts, rotations, and scaling. For digital differential rectification, each pixel is separately transferred form the original image to the resulting image using the indirect approach (Equation 1).

A digital elevation model (DEM) is needed to correct for relief displacements in the image. It is assumed that the DEM is stored in the same format as the digital image, i.e., as a raster of digital values, which in this case expresses elevations instead of densities and are 16 bits long (compared to 8 bits of gray value). The DEM has to be related to the map projection of the resulting orthophoto.

In the digital case, the objective of rectification is the assignment of gray values from the aerial photograph or satellite image to each pixel of the DEM. After the rectification, both elevation and density are stored at the same (x, y) location. This is basically similar to the concept of groundels which has been suggested by Helava in connection with the automatic derivation of DEMs (Helava, 1988).

To determine gray values, we follow this procedure: The threedimensional coordinates (X, Y, Z) defined by a DEM pixel are transformed into the image by the collinearity equations (Equation 5). At the image position (x, y) the gray-value is interpolated by one of the resampling methods. The density is stored at the X, Y location of the digital orthophoto, which is equal to the position of the DEM point (Figure 4).

$$x = x_{p} - c \frac{r_{11}(X - X_{0}) + r_{21}(Y - Y_{0}) + r_{31}(Z - Z_{0})}{r_{13}(X - X_{0}) + r_{23}(Y - Y_{0}) + r_{33}(Z - Z_{0})} = f_{x}(x', y')$$

$$y = y_{p} - c \frac{r_{12}(X - X_{0}) + r_{22}(Y - Y_{0}) + r_{32}(Z - Z_{0})}{r_{13}(X - X_{0}) + r_{23}(Y - Y_{0}) + r_{33}(Z - Z_{0})} = f_{y}(x', y')$$
(5)

where x', y' are equivalent to the map coordinates X, Y.

To perform this transformation, the following parameters have to be available:

- the interior orientation of the camera: X_p, y_p, c,
 the exterior orientation of the camera: X_o, Y_o, Z_o perspective center, R a rotation matrix composed of ω , φ , κ -rotation angles,
- the pixel-spacing of the digital image in camera units: px, py (mm),
- the cell-size of the DEM pixels in ground units: gx, gy (m),
- · the reference coordinates of one DEM pixel in the given mapprojection (usually the left upper corner of the DEM file).

If SPOT images are to be used, the mathematical model has to be modified to line-perspective geometry (Equation 6). That is,

$$x_{i} = -c \frac{r'_{11}(X - X_{io}) + r'_{21}(Y - Y_{io}) + r'_{31}(Z - Z_{io})}{r'_{13}(X - X_{io}) + r'_{23}(Y - Y_{io}) + r'_{33}(Z - Z_{io})} = f_{x}(x', y')$$

$$0 = -c \frac{r'_{12}(X - X_{io}) + r'_{22}(Y - Y_{io}) + r'_{32}(Z - Z_{io})}{r'_{13}(X - X_{io}) + r'_{23}(Y - Y_{io}) + r'_{33}(Z - Z_{io})} = f_{y}(x', y')$$
(6)

where x, is the coordinate in scan-line i, orthogonal to the di-

rection of travel, and x_p , $y_p = 0$. Basically, the orientation parameters of each scan line are different. However, due to the smooth motion of the satellite along the orbit and the relatively short time it takes to capture a full scene, neighboring perspective centers are highly correlated. For most purposes the change of the exterior orientation

parameters can be approximated by linear functions (Equation 7). That is,

$$X_{oi} = X_{0} + k_{x}y_{i}$$

$$Y_{oi} = Y_{0} + k_{y}y_{i}$$

$$Z_{oi} = Z_{0} + k_{z}y_{i}$$

$$\omega_{i} = \omega + k_{\omega}y_{i}$$

$$\varphi_{i} = \varphi + k_{\varphi}y_{i}$$

$$\kappa_{i} = \kappa + k_{x}y_{i}$$
(7)

where

X_{oi} - κ_i	exterior orientation parameters of line i,	
X_0 - κ	exterior orientation parameters of the	
	center-line of the scene,	
$k_X - k_\kappa$	linear coefficients of the exterior orienta-	
	tion, and	
y_i	line number (coordinate component in the	
1000 I	flying direction).	

In this case, there are 12 parameters of exterior orientation for each image compared to six for regular aerial photos. Additionally, the Earth curvature has to be considered with satellite images and small scale aerial photographs. The reference of the orthophoto is a map projection related to the Earth ellipsoid (e.g., UTM, Lambert's Conformal). The coordinate triplet (X, Y, Z) of the collinearity equations, however, is related to a tangential plane touching the ellipsoid at the center of the image (Figure 5). Due to the large coverage of a satellite scene (60 by 60 km for SPOT), the differences of the map and plane coordinates cannot be neglected. The largest influence appears at the elevations (Z-component); at the edge of the satellite image the height-difference can be more than 70 m, as compared to a vertical accuracy of about 10 to 20 m. To transform a pixel location (x', y') of the orthophoto to the original image, one has to go from the map coordinates (X_M, Y_M, Z_M) with $X_M = x', Y_M = y'$ to the tangential plane coordinates (X, Y, Z), and from there (X, Y, Z) to the satellite scene (x_i, y_i) utilizing Equations

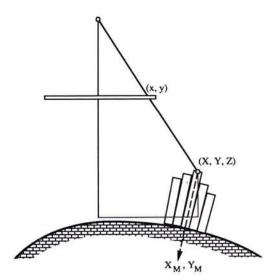


Fig. 5. Correction of the Earth curvature in satellite images and small scale aerial photographs. The map coordinates $X_{\rm M}$, $Y_{\rm M}$ have to be transformed into a Cartesian system (X, Y, Z).

IMPLEMENTATION CONSIDERATIONS

As every pixel of the orthophoto has to be processed separately, computation time is a major concern. It can be largely reduced by optimizing the image transformation. Multiplications must be avoided and substituted by simple additions. This is done by taking advantage of the incrementation of pixel locations at the resulting image (orthophoto).

As an example, the projective transform (Equation 3) is used; practically, this technique can be applied to any transformation. With the indirect approach we take a pixel of the result, compute its corresponding position in the original image, and interpolate a gray value. Then we proceed to the next pixel by adding 1 to the x-coordinate. The left, upper corner of the orthophoto has pixel-coordinates x' = y' = 0; the transformed values are given by Equation 8; that is,

 $x_{0,0} = a_3/1$ (8) $y_{0,0} = b_3/1$

Then the next pixel of the same line (x' = 1, y' = 0) is transformed by adding a_1 , b_1 to the numerators and c_1 to the denominator of Equation 8; no multiplication is necessary (Equation 9). That is,

$$x_{1,0} = \frac{a_1 + a_3}{c_1 + 1}$$

$$y_{1,0} = \frac{b_1 + b_3}{c_1 + 1}$$
(9)

For each following pixel of this line the same three additions are done to get the location in the original image. When we switch the line (y'=1), we start by adding a_2 , b_2 and c_2 to the numerators and denominators of $x_{0,0}$ and $y_{0,0}$, respectively. Thus, the transformation of all pixels of this orthophoto can be performed without a single multiplication; of course, the division of the projective transform is still necessary.

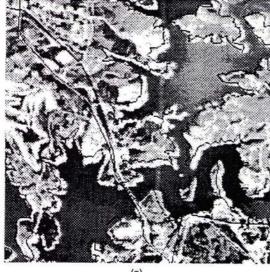
COMPARISON OF METHODS

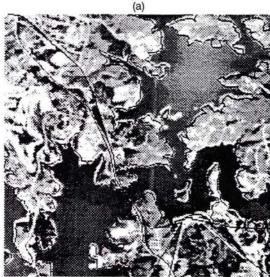
The three methods have been implemented on a workstation. The whole orthophoto production, consisting of the orientation of the images and the rectification, was done digitally. First, the image coordinates of control points were measured on a monitor employing a cursor and a mouse. The reference points on the ground were digitized from an existing map. The parameters of each transformation were computed from these measurements. The residuals of image coordinates after the adjustment were displayed as vectors on the screen to evaluate the quality of the points and to eliminate blunders. Afterwards, the transformation parameters were used in the rectification programs. For differential rectification a DEM was available with the same cell size as the orthophoto. It was captured from the original aerial photos on an analytical plotter.

Figure 6 shows different rectifications of a part of an aerial photograph, which was digitized with an EIKONIX camera at a format of 1024 by 1024 pixels. Image (a) was done by polynomial rectification, image (b) by projective transformation, and image (c) by differential rectification. The streets and lakeshores were digitized manually from an existing map and overlaid over the orthophotos. The polynomial approach shows the largest displacements, while both the projective transformation and differential rectification fit very well to the digital map. It seems that the ground plane of the projective transform is tied to the lakes, thus showing almost no deviations from differential rectification.

CORRECTION OF DISTORTIONS

The accuracy of differential rectification can be further improved by extending the analytical model. Parameters ($\triangle x$, $\triangle y$)





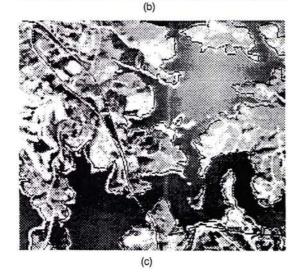


Fig. 6. Comparison of different rectifications of the same area. The images have been overlaid by a digital map. While the polynomial transform (a) caused very large displacements, projective (b) and differential (c) rectifications are very similar, as the relief displacement has only a small influence due to the large flying height.

were added to the collinearity transformation to correct for all kinds of distortions of the digital image. This is especially important if CCD and video cameras are used for imaging, as their lenses are of lower quality than those of metric cameras and show considerable distortions. The coefficients implemented are correcting symmetric and asymmetric radial distortions, tangential distortions, as well as affinities (Equation 10) (Konecny, et al., 1984). They can be added to the image coordinates of the projective transformation and the collinearity equations. They are obsolete for polynomial rectification, as most of the coefficients are correlated with the parameters of the polynomial.

$$\Delta x = a_1 x (r - r_0) + a_2 x (r^3 - r_0^3) + a_3 x (r^5 - r_0^5) + a_4 x (r^7 - r_0^7) + a_5 x \cos 2\alpha + a_6 x \sin 2\alpha - a_7 y \cos 2\alpha - a_8 y \sin 2\alpha - a_9 x$$

$$\Delta y = a_1 y (r - r_0) + a_2 y (r^3 - r_0^3) + a_3 y (r^5 - r_0^5) + a_4 y (r^7 - r_0^7) + a_5 y \cos 2\alpha + a_6 y \sin 2\alpha + a_7 x \cos 2\alpha + a_8 x \sin 2\alpha - a_9 y + a_{10} x$$

$$(10)$$

where

 $a_1 \dots a_{10}$ are additional parameters computed by the bundle adjustment,

r, r_0 are the radial distances from the principal point, and α is the angle between the x-axis (horizontal) and the direction to the image point.

To prove the effectiveness of these corrections, a video image of a two-dimensional test field was rectified with and without the added parameters (Figure 7). The target locations were measured by least-squares template matching in both rectifications and compared to the correct point coordinates. The results are show in Table 2. The mean accuracy of the points in the rectified images improved by a factor of three to about half a pixel in both directions.

CONCLUSIONS

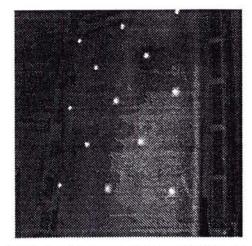
Digital rectification methods are gaining importance due to the fact that images must be combined with digital maps in GISs. The higher the resolution of the images, the more sophisticated the rectification method should be. All possible distortions caused by the imaging sensor can be corrected during the rectification to obtain best results. The comparison of three different approaches commonly used in image processing systems proved that a differential rectification that corrects both for relief displacements and camera distortions yields the best results, and should be used both for aerial and satellite images.

This method not only generates orthophotos, but it actually finds the gray value of each DEM point and stores it at the same location. Thus, a digital surface model is created which contains elevation, density, and maybe other attributes (such as exposure, slope, temperature, etc.) at each position. These different layers of raster data are given in a geo-referenced, orthogonal projection, and can easily be overlaid by vector data stored in the GIS.

Moreover, both orthophoto and DEM can be created at the same time by digital image matching. During the matching procedure the gray value which has to be assigned to a given location is available twice (in the left and in the right image). Therefore, the mean value could be used. One might even develop more sophisticated methods that determine which of the gray values yields a better result, and thus solve the problem of occlusions.

REFERENCES

Baehr, H. P., 1985. Digitale Bildverarbeitung - Anwendung in der Photogrammetrie. Herbert Wichmann Verlag, Karlsruhe.



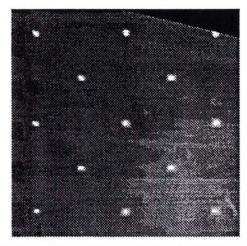


Fig. 7. The digital image of a testfield (left) has been rectified (right) with and without corrections for distortions. The improvement has been quantified by measuring the image coordinates of the two rectifications by least-squares matching.

Table 2. Standard Deviation of Point Positions in Digital Images Rectified With and Without Applying Additional Parameters

Additional Parameters	Standard Deviation (RMS Error)	
	x	y
Without	1.65	1.73 [pixels]
With	0.56	0.63 [pixels]

Helava, U. V., 1988. Object-Space Least-Squares Correlation. Photogrammetric Engineering & Remote Sensing, Vol. 54, No. 6, pp. 711– 714.

Hood, J., L. Ladner, and R. Champion, 1989. Image Processing Techniques for Digital Orthophotoquad Production. *Photogrammetric Engineering & Remote Sensing*, Vol. 55, No. 9, pp. 1323–1329.

Jansa, J., 1983. Rektifizierung von Multispektral-Scanneraufnahmen. Geowissenchaftliche Mitteilungen der TU Wien, Heft 24.

Keating, T. J., and D. R. Boston, 1979. Digital Orthophoto Production Using Scanning Microdensitometers. Photogrammetric Engineering & Remote Sensing, Vol. 45, No. 6, pp. 735–740.

Konecny, G., and P. Lehmann, 1984. Photogrammetrie. Walter de Gruyter, Berlin, New York.

Konecny, G., 1979. Methods and Possibilities for Digital Differential Rectification. Photogrammetric Engineering & Remote Sensing, Vol. 45, No. 6, pp. 727–734.

Mueller, W., and H. Sauleda, 1988. Orthophoto Production in the new ContextMAPPER System. *International Archives of Photogrammetry* and Remote Sensing., Vol. 27, B 9, p. II-224 - II-231.

(Received 8 February 1991; accepted 10 April 1991; revised 13 May 1991)

1st Australian Conference on Mapping & Charting "Mapping for A Green Future"

Adelaide, South Australia 14-17 September 1992

The Australian Institute of Cartographers is hosting a conference focusing on the role of cartography in monitoring and managing the nation's resources. Conference themes include:

Remote Sensing & Geographic Data for Environmental Management • Mapping Management & Professional Practice • Future Directions in Education & Training • Development of Mapping & Charting • International Mapping Opportunities • Thematic Cartography • GIS/LIS Development

For further information, contact:

Conference Secretariat
1st Austrailian Conference on Mapping and Charting
GPP Box 1922, Adelaide
SOUTH AUSTRALIA