DEM Determination from SPOT

C. Angleraud
K. Becek, and J.C.Trinder
School of Surveying, University of NSW
P.O. Box 1, Kensington, 2033 Sydney, Australia

ABSTRACT

Software developed for computation of DEM's is based on a satellite orientation model which varies according to polynomials as a function of spacecraft position. The determination of digital elevation models (DEM) from the digital data incorporates this satellite model for the computation of ground coordinates from the two overlapping digital images. The process of computing the matched image points has been divided into two steps. The first is based on the extraction of features on the two images, while the second is based on grey level matching using the pixel intensity values within windows on the two images. The process is based on a triangulated network which progressively densifies points within each triangle. Results of accuracy tests using this method are presented in this paper.

KEY WORDS: DEM, Image Matching, Image Processing, 3D.

1. INTRODUCTION

SPOT satellite images are the first data that can be considered suitable for mapping at small and medium scales. Tests in several locations around the world indicate that SPOT images are well suited to mapping and map revision at 1:50,000 and 1:100,000 (Dugan and Dowman 1988, Murray and Farrow 1988). The suitability of the images for mapping depends on the two primary requirements which must be satisfied; they are, the geometric accuracy of features extracted from the images, and the interpretation of adequate features from the images for specific map scales. Generally it has been found that the geometric accuracy of the data extracted from the images can satisfy larger map scales than can be satisfied by the number of features interpreted from the data. Mapping and the determination of digital elevation models (DEM) from SPOT images can be based on analogue images produced from the digital data using analytical stereoplotters (normally level 1A or IAP images are preferred), or the digital images themselves. Both types of data have been investigated by the authors but this paper will concentrate on the development of software for the computation of DEM's.

The choice of whether to observe analogue images on an analytical stereoplotter, for the derivation of DEM's or compute DEM's from digital data, ultimately rests on the assessment of the time and cost of manual observations, versus the cost of the computations from digital data. Based on an observation rate for a human observer, of one height observation every 5 to 10 seconds, the time required to observe a grid of points over a complete SPOT stereo-pair, even at an interval of 250m, can amount to several weeks. The cost of such an operation is likely to be very high, while the sheer boredom for the operator in observing many thousands of points is likely to affect the accuracy. Software to carry out this process on the digital data requires considerable time to develop because of its complexity and will most likely require some interaction from the operator in difficult terrain cases. However, once the software has been developed to a high level of automation, it will not require continuous monitoring from the operator. The cost of the process is therefore likely to be significantly less than if the DEM is derived by manual means. A test of the accuracy of the manual observations of a Wild-Leitz BC2 compared with 2m contours of the area indicated results of approximately 5m. The same DEM's will be compared with that derived by the software using the digital images.

1.1 SATELLITE MODEL

The satellite model that has been used for the mapping and computation of DEM's is based on an image to ground relationship shown in equations (1) and (2), and as demonstrated in Figure 1. Because of the nature of the linear array or push-broom scanner which is used to acquire the SPOT images, a separate set of collinearity equations must be written for each scan-line.

$$F_{xj} = x_j + f\,\frac{m_{11}(X_j - X_k^c) + m_{12}(Y_j - Y_k^c) + m_{13}(Z_j - Z_k^c)}{m_{31}(X_j - X_k^c) + m_{32}(Y_j - Y_k^c) + m_{33}(Z_j - Z_k^c)} = 0$$

$$F_{yj} = f\,\frac{m_{21}(X_j - X_k^c) + m_{22}(Y_j - Y_k^c) + m_{23}(Z_j - Z_k^c)}{m_{31}(X_j - X_k^c) + m_{32}(Y_j - Y_k^c) + m_{33}(Z_j - Z_k^c)} = 0$$

Equations (1) and (2)

where X_k^c, Y_k^c, Z_k^c are the coordinates of the perspective centre for line k of the image

f is the effective focal length

$m_{11}, m_{12}, ..., m_{33}$ are the elements of the rotation matrix of the sensor for line k

X_j, Y_j, Z_j are the object coordinates of point j.

Figure 1 shows the image space/object space geometry for a point 'j' for a linear array image. Taking the image x-axis to be at right-angles to the satellite path, the projective relationship exists only in the x-z plane of the image space. Thus, the resulting collinearity equations for a line 'k' are given in equations 1 and 2. A set of these equations is written for each line of the image with its own perspective centre and attitude. The parameters of the perspective centre and attitude are modelled by polynomials in terms of the y-image coordinate, with the position parameters following second order polynomials and the tilts following first order polynomials. The solution of the orientation of two overlapping SPOT images involves a total of 30 unknowns. Ephemeris data derived from the image header file can also be used as input for the modelling of the polynomials. This

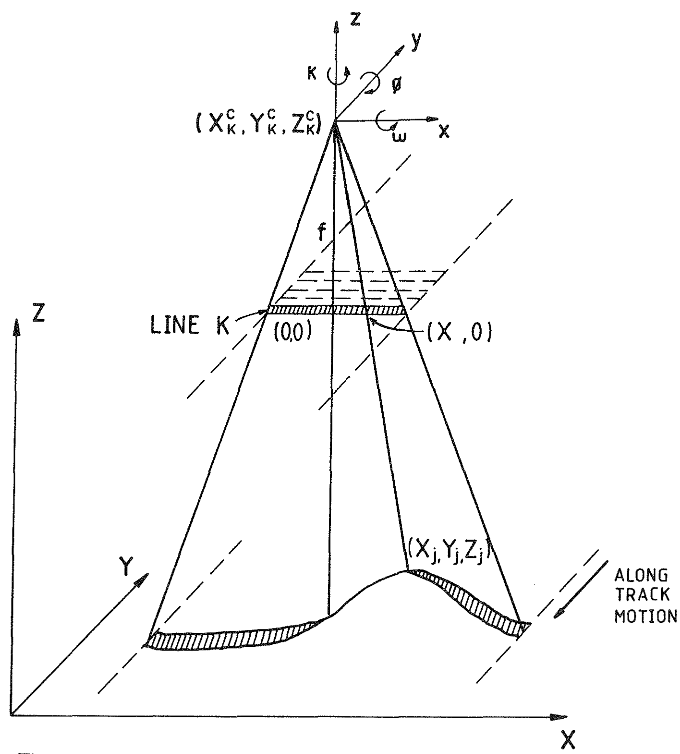

Figure 1. Collinearity Relationship for Linear Array Imagery

model for the satellite model has been used in the SATMAP software used for mapping from analogue images on the Wild-Leitz analytical stereoplotter and also has been implemented in the software used for the computation of DEM's.

2. METHODOLOGY

The aim of the software described in this paper is to automatically compute a DEM over an area covered by two overlapping SPOT images. To achieve this goal, the software must be able to locate and match the corresponding features on the left and right images. This matching process must be precise (around one third of the pixel size) and reliable.

To meet those two requirements, a methodology has been adopted which allows the software to take advantage of as much of the information available from different sources as possible. For example, it should be possible to use the height information contained in a set of Ground Control Points (GCP) as well as the set derived by feature based operators, for isolated or linear features. In order to bring all this information together into a homogeneous environment, the DEM software has been based on Delaunay triangulation principles.

Basically, the process is as follows:

- Manually extract about 10 - 15 GCPs for the computation of the orientation parameters of the two images.
- Integrate these GCPs into a Delaunay triangulation.
- Determine new matching points using the height information provided by the triangulation, to predict the areas that should be matched. Different approaches can be adopted for this process, the most common one being the identification of "interest" points by a feature operator such as the Moravec operator, followed by a photogrammetric least square matching of two local windows.
- The process is repeated until the densification of the network is judged sufficient for the grid size required.
- At any stage the DEM can be checked manually, either by the display of perspective views or by superimposition of two (left and right) ortho-images.

2.1 JUSTIFICATION FOR THIS APPROACH

As the integration of the multiple sources of information is one of the major goals of the approach, a method which provides speed, simplicity and versatility is required. The Delaunay triangulation fulfils these requirements. As the DEM is being built, the matching information provided by the different modules can be integrated, regardless of its source or position on the ground. When a new point is found, it is added to the existing network. If, subsequently, it appears that it was an incorrect match, it can either be corrected or eliminated. Moreover, it is possible to interpolate the height very quickly at any location based on the triangles, thus giving a fast method of predicting the location of a feature on the right image which corresponds with a feature on the left image.
It is also possible to use an existing DEM of any grid spacing to guide the matching process, by integrating height information from a separate triangulation.

2.2 SYSTEM OPERATION-PROGRAM STRUCTURE

The program has been designed for future operational use in a digital workstation environment, based on a modern, window type interface and requires minimal intervention by the user.

When the program is initiated, three windows are displayed onto the screen. The largest one contains an overview of the left image, the second window a series of icons which command different actions, while the third is a multi-purpose window, basically used to display some enlargements of the image in different fashions. All the parameters concerning the file names, the window sizes of the different operators, the

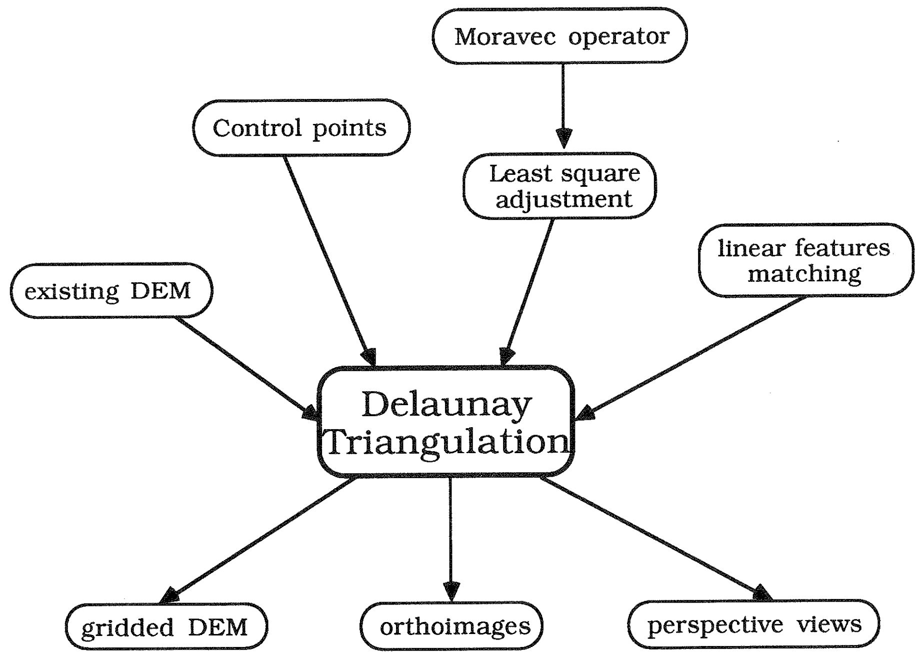

Figure 2.

position of the icons and the windows on the screen are gathered in one configuration file whose syntax is understandable by the operator : *keyword = value*.

The GCPs are measured manually on a screen, based on enlargements of small windows of the two images and the orientation subsequently derived, according to the formulae described in Section 1.1. The GCPs are then triangulated, and the user can display this triangulation as an overlay. Numerous possibilities are then available, e.g.: zoom in on any area at any scale; or display any area in stereo in two separate subwindows at true resolution and at five times the true resolution to precisely identify features (road crossings, river forks, buildings etc.). The user can thus select an area where network densification is required and start the automatic matching process.

There are several procedures to allow the operator to check and control the generation of the DEM. Firstly, the user can display some areas in stereo and check if the points added are truly matching. He/she can, as well, display a perspective view to test the shape of the model. Left and right ortho-images can also be generated so that the user can check visually if they correspond, and if not, identify the areas which have been incorrectly matched. It is then possible to alter some of the points of the triangulation or even to suppress them.

As all these operations are highly interactive it is vital to have a system which responds quickly to any user request. The program must be carefully designed to avoid serious operational problems for the user.

3. IMAGE MATCHING

The image matching is based on a sequential process of feature matching followed by area based matching. Each of the steps will be described below.

3.1 SELECTION AND MATCHING OF FEATURES

The feature matching of SPOT images is based on point features or 'interest points' after Moravec (1977). The extraction of the interest points from images, according to Moravec method is undertaken in two steps:

- For every pixel in both the left and right image a value of a so-called interest operator is calculated.
- Thresholds of the values of the interest operator for both images are chosen, and those pixels above threshold in both images are qualified as interest points.

In the case of the Moravec method, the interest operator is the minimum of the slopes in intensity derived in four directions. The evaluation of the interest operator commonly uses a 5x5 window around each pixel which is a very time consuming task. In addition, the determination of thresholds for the interest operator for whole images can often be difficult, especially in the case of highly unhomogeneous images.

Matching of interest points selected in both images may typically be based on criteria such as those described by Barnard and Thompson (1981). That is, the matched interest points must have similar characteristics of discreteness, similarity, and consistency. Discreteness is determined by the interest operator. Similarity is determined by the similarity of the intensity patterns of the potentially matching interest points on both images. Based on this test, the decision is made as to which of all possible combinations of interest points on the two images may potentially match by virtue of the intensity distributions around them. The final test of consistency aims to eliminate erroneous matchings by testing the consistency of the geometry of the sets of possible matching points on the two images. This can be done typically by an affine transformation of the positions of potential matched points in each image.

In order to avoid the need to perform the very time consuming task of generating the so-called Moravec image, the method is modified so that the interest points are selected in a set of small sub-areas distributed over an image. A sub-area is typically 30x30 pixels in size and its centre is positioned in the left image by arbitrarily choosing points such as nodes of a regular grid or of the triangular network. The corresponding centres of the sub-areas in the right image are predicted using a priori information about the stereo-model and the topography of the terrain itself (roughness of the terrain and average elevation). Compared with the original Moravec method this approach significantly reduces the time needed for generating the Moravec image. As a result a discrete Moravec image is derived instead of a continuous image.

In every sub-area of both images only one interest point is selected, the one with the maximum value of the interest operator. The implication of this step is that there is no need to set a threshold for the value of the interest operator, as is done in the second step of the original Moravec method. At this stage it is assumed that an interest point selected in a sub-area of the left image should match the interest point similarly derived in a corresponding sub-area of the right image. Investigations have shown that about 60 - 70% interest points are matched correctly.

In order to assess whether a pair of interest points will match, a test of similarity has been introduced. For every interest point selected in both images, three parameters are calculated which characterise the intensity patterns around them. These parameters are the ratios between mean gradient slopes in the intensity values derived for the four principal directions for a particular interest point. The mathematical formulas for the similarity test are as follows.

Assume a window sized 9 x 9 pixel is placed on an interest point. Let the intensity gradients of neighbouring pixels within the four principal directions (d= 0,..,3)

$$g_i = I_{i+1} - I_i , \tag{3}$$

where: $i = 0, 1, 2.$

and I is the intensity value.

The mean gradient in intensity values is defined by:

$$S^d = E(g_d) \tag{4}$$

d = 0,...,3, where E(.) is the operator of mean value. The ratio of mean gradients is therefore defined by:

$$R_d = \frac{S^{d+1}}{S^d} \tag{5}$$

where $d = 0,1,2.$

The rejection threshold applied for the measure of similarity of interest points on 2 corresponding windows is:

$$R = \frac{1}{n}\sum_{j=0}^{n}\sqrt{\frac{\sum_{i=0}^{2}(R_i^L - R_i^R)^2}{3}} \tag{6}$$

where n is the number of interest points and R^L, R^R are the corresponding ratios for the pair of interest points being matched in the left and right images.

If for a pair of interest points the average difference in ratios is greater then $R + 2.0^* \sigma_R$ then matching of the pair is considered as incorrect. In the sub-areas of the right image, for which the matching is incorrect, a further set of four points is selected. These are chosen according to descending values of the interest operator for pixels in those sub-areas. On this set of additional interest points the similarity test is also performed. Matched points are selected accordingly to the smallest difference from the average of the ratios of the gradients.

The evaluation of performance of these similarity tests of interest points have hsown that the number of correctly matched interest points has increased to a level of 90 - 95%. The advantage of this similarity test is that the ratio of the mean gradients are calculated based on previously derived values of gradients in the intensity value, hence eliminating recomputation.

Linear feature based matching will also represent a component of this software, creating a lattice of matched features over the image. The techniques to be used for this process, which has not yet been incorporated into the software, will be described by Butler (1992).

3.2 AREA BASED MATCHING.

Grey level or intensity based matching is carried out by the least squares method, based on 6 affine transformation parameters, as has been used by many photogrammetrists to achieve high precision matching e.g. Gruen and Baltsavias (1985), Rosenholm (1987).

The least squares matching method has the advantage that by its very nature, distortions in geometry can be corrected through the resampling process, and in addition, it provides information on the quality of the match through the weighted sum of squares of the residuals at the pixels. Other forms of error detection available in the least squares method can also be used. Accuracies of the method are typically 0.3 pixel, Gruen and Baltsavias (1985), Rosenholm (1987).

In this project the least squares matching is performed on a window of 21 x 21 pixels. The thresholds for the shift and rotation terms in the solution are set at 0.1. Average number of iterations needed to satisfy the threshold conditions is 3 to 4. Of course, any matching computation which does not converge in the prescribed number of iterations is discarded.

An essential element of this computation are the procedures for checking the accuracy of the matching and the computed elevations. This is divided into 3 stages:

* checks in the image space. The rotation parameters in the least squares solution should be similar within certain regions of the overlap area of two images, and variations in these parameters will be largely due to the effects of parallaxes caused by variations in terrain elevations. The standard deviations of the rotation parameters derived from the matches within each region are compiled during the computation. Any parameter which deviates from the mean by more than a certain factor times the standard deviation will be discarded as an erroneous match. Further, a y-parallax between the computed match points greater than 2 pixels will cause the point to be discarded.

* those which check the distribution of the elevations. The RMS variations in the elevations within a certain region will indicate the characteristic shape of the terrain. Therefore, outstanding elevations in a certain region of the terrain will be identified as those which are greater than the average terrain shape, by a set factor of the RMS variations. These points will normally be discarded unless further tests indicate that the terrain variations are indeed due to a marked change in elevation. In addition, when a grid of points is being computed, the distance between points is also used as a check parameter.

* visual observations on the computed data. This test could involve viewing the stereomodel in a digital workstation where available, but at present, this test will be based on the observation of overlaid ortho-images.

4. PERFORMANCE OF THE SOFTWARE

There are many parameters to be taken into account in assessing the performance of the package.

- the user interface;

- speed;
- accuracy of the generated DEM.

Regarding the user interface, the use of icons, windows, point and click selections, makes the whole package easier to use. Response time is also important. This aspect has been borne in mind for any procedure that is run interactively. For example, two different drawing algorithms co-exist for drawing a triangulation network, one for the overview, and another for enlargements which explores and draws the network locally. The way the windows are refreshed has also been carefully considered.

The most time consuming part of the software is by far the least square matching package. For a window size of 21 x 21 pixels and for a mean number of 4 to 5 iterations, there are 50 000 double precision floating point multiplications. It has been estimated that with a 20 Mflops processor it is possible to compute a DEM or a complete SPOT stereo pair in 6 to 7 hours. The software has therefore recently been ported to a HP730 RISC processor.

It has been necessary to pay attention to the way the information flows through the different loops. Some examples are: limiting the number of conversions between bytes (pixel intensity values) and floating point numbers; suppressing unnecessary multiplications (during iterative loops it is a lot faster to add a step than to multiply the index of the loop by the given step); to optimize file transfers, image files should be rewritten in blocks, each block being a small window of the image with size matching that of the smallest block transfer provided by the operating system. This facility can improve the performance of the system by up to 60%.

Initial accuracy tests have been made by comparing the computed elevations with a DEM that has been observed on the Wild-Leitz BC2 analytical stereoplotter with the SATMAP software. 55,000 points at 250m interval were manually observed over the Sydney region on an image pair with a B/H ratio of 1.0, with an accuracy, tested against maps at a scale of 1:4,000, of 5 to 6m. The location of points in this DEM did not match those derived by computation which was made at an interval of approximately 120 m. A weighted average elevation was derived at the coordinate positions of the manually observed DEM from the nearest 3 computed DEM points. This method of interpolation may lead to slight errors in interpolation, but since most of the terrain elevations vary gradually in the Sydney area, the errors are likely to be small. A total of 9000 points in the computed DEM were used for the comparison, corresponding to 1700 points in the manually observed DEM. The RMS variation between the two DEM's was then computed but large residuals greater than 27m (3 x RMS) were discarded in this computation. The total number of discarded residuals was about 5% of the total number computed. The RMS derived was 9m with an average value of 1 m. This accuracy estimate is influenced by the accuracy of the observed DEM and the effects of the interpolation. If the manually observed DEM is assumed to have an accuracy of 5m, the accuracy of the computed DEM is of the order of 7m. Further tests will be carried out with new scenes currently under investigation. These results compare with those published by Theodossiou and Dowman (1990) which indicate accuracies varying from 6m to 25m depending on the terrain slope, with considerable variation in the performance for a given terrain slope.

The stereopair tested were recorded with a 6 weeks time difference between the two images and some specularity problems. Furthermore, the Sydney area is a difficult environment because the density of buildings and roads can result in brightness values, while forest areas appear very dark. As well, the images were taken in 1986 when the early column noise of SPOT had not yet been corrected. Despite these problems, the Moravec-Least Square approach has been robust. But even if more than 95% of the points generated are correct matches, some gross non-matching errors still occur. Methods are currently being developed for the identification of some parameters derived from the Moravec operator and the least square adjustments that would allow the software to automatically discard these errors. This will involve a learning process to calculate the influences of the different parameters and is a first step towards the use of artificial intelligence techniques in the domain of automatic DEM computation. The inclusion of the matched linear features into the software will also be an important source of feature information for determining matched points.

References

Butler, N. (1992). "Linear Based Matching of Stereo SPOT Satellite Images". Paper submitted to ISPRS Congress,

Barnard, S.T. and Thompson, W.B. (1981). "Disparity Analysis of Images. IEEE Trans. PAM I - 2, pp. 330 -

Gruen A.. & Baltsavias E.P. (1985). Adaptive Least Squares Correlation with Geometric Constraints. SPIE, Vol. 595,

Rosenholm, D. (1987). "Multipoint Matching Using Least Squares Technique for Evaluation of Three-Dimensional Models". Photogram. Eng. & Rem. Sens. Vol. 53, pp.

Sloan, S.W., (1986). "A Fast Algorithm for Constructing Delaunay Triangulation in the Plane", Report from the Department of Civil Engineering and Surveying, The

Theodossiou, E.I. and I.J. Dowman 1990 'Heighting Accuracy of SPOT' Photogram. Eng. & Rem. Sens. Vol. 56 pp 1643-1649.

Trinder, J.C., Donnelly, B.E. & Kwok, L.K. (1988). "SPOT Mapping Software for Wild Aviolyt BC2 Analytical Plotter". International Archives of Photo. & Rem. Sens., Vol. 27, Pt B4, pp. 412 - 421.