MULTIPLE-IMAGE SAR SHAPE FROM SHADING

J. Thomas, W. Kober, F. Leberl
VEXCEL Corp.
2905 Wilderness Pl.
Boulder, CO 80301, USA
Tel: 303/444-0094
FAX: 303/444-0470

ABSTRACT

The purpose of the present effort is to explore the combination of techniques for shape from shading with stereo-radargrammetry to produce terrain surface models using multiple SAR imagery.

The use of local variations in pixel shading as an indicator of terrain slope changes represents an opportunity for increasing the accuracy of terrain mapping over that which is available from stereo-radargrammetry alone. Shape from shading can potentially provide a relative changes in height at each pixel. This leads to a denser set of height measurements and a more faithful rendition of the local terrain shapes. However, some additional type of boundary values or terrain low-frequency information is required. These can be obtained from stereo or from altimeter measurements.

Two essentially different types of approaches are considered. Preliminary results using real SAR imagery are encouraging.

1. INTRODUCTION

Shape from shading requires the formulation of the reflectance map, which is an analytic description of the relationship between radiometric image values I (x,y), surface coordinates, and image sensing vectors. We will assume the following parameters are the determinants of the reflectance map, R [Frankot,Chellapa,87]:

(1) $I(x,y) = R(Z_x, Z_y, b, l, n)$

where:
Z_x, Z_y = slopes in x and y directions
b = illumination vector
l = boresight vector
n = albedo

Note that for radar, b = l

The image gray values are then the discretized levels between the minimum and maximum radiometric values.

An exact functional relationship among these variables is intractable to derive analytically, and in the past has been the subject of empirical
studies and some simplified modeling of the scattering mechanisms involved.

However, in many shape from shading studies, for example [Wildey,86], the reflectance map is simply the product of a constant scattering cross section coefficient, dependent only on an assumed constant albedo value, and the cosine of the local incidence angle, ie. the angle between the illumination vector and the local terrain surface normal vector. In the single image case, there is a fundamental ambiguity between changes in albedo and slope when a change in shading is encountered.

Such a model assumes that the scattering cross section coefficient so, does itself not vary with local incidence angle. In order to simplify the computations involved, this is the model that has been assumed in the present study. More complex models of the reflectance map will be used in follow-on efforts.

A major difficulty with single image approaches to shape from shading is the assumption of uniform albedo. Another difficulty with shape from shading is that the problem is mathematically underdetermined without boundary conditions. In general, the shading information from single pixels merely implies a cone constraint for the local surface normal. Continuity assumptions connect the estimated directions for neighboring pixels, and prevent independent solutions for each separate pixel.

Another constraint involves the notion of "integrability". Analytically, this condition implies that:

(2) $Z_{xy}(x,y) = Z_{yx}(x,y)$

where z(x,y) is the analytical expression for the height.

This analytic condition is just the intuitive condition that heights can be integrated along any path, since these values are independent of the path of integration. In practice, enforcing this condition also acts as a smoothing process on the computed terrain surface.

Still another constraint is "regularization". This term effectively limits the amount of allowable oscillation in the computed terrain surface.

Two essentially different approaches for reconstructing the surface are discussed in this papaer. Approach #1 consists of two variations which involve a multiple image generalization of the method in [Frankot,Chellapa,87]. The latter approach formulates the reconstruction problem as the minimization of a cost function which contains two types of terms.

The first type of term is a measure of the difference between the pixel gray values of an actual image versus the values predicted using the current estimated terrain model. Because of the mathematically ill-posed nature of such inverse problems [Baltes,80], solutions which minimize this term will lead to very oscillatory terrain estimates.

Therefore, the second type of term is a regularization term which acts as a penalty function to limit the amount of terrain oscillations. In this way, solutions to the combined metric represent a compromise between faithful image prediction and terrain slope variations. The reconstruction problem is thus formulated as a calculus of variations problem whose solution is obtained using the Euler-Lagrange equations.

Such methods tend to be robust with respect to noise because they tend to distribute the effects of noise, rather than accumulating them as in an integration algorithm.

However, most of the earlier work done using the variational approach tended to produce solutions that did not satisfy the condition of integrability. One approach [Horn,86] used a penalty function to drive the iterated solution toward integrable solutions. In practice, this meant that the computed solution usually was only "close" to being integrable. A step in the right direction was taken in [Frankot,Chellapa,87] wherein the latest estimated solution was projected onto an integrable subspace of solutions on every iteration.

However, both variations of Approach #1 require the mutual registration of multiple images. Because such registrations generally involve residual errors, an iteration between registration and terrain reconstruction is generally required.

Approach #2 for multiple imagery involves a set of simultaneous equations which relate slope-induced shading and height measurements to the Fourier- series coefficients of the terrain model.

2. ALGORITHMS FOR SHAPE RECONSTRUCTION

2.1 General

Section 2.2 discusses the two variations of approach #1, the N-image generalization of the method in [Frankot, Chellapa,87]. A very brief description of approach #2 is given in section 2.3.

2.2 Approach #1: Generalizations of Frankot's and Chellapa's Method

The two generalizations of Frankot's and Chellapa's method involve generalizing the cost function or combining separately derived models using projection in Fourier space. These approaches follow in the next two sections. A discussion of the incorporation of spot heights derived from stereo into the reconstruction is described in section 2.2.3.

2.2.1 Variation A: Generalized Cost Function

One approach to an N image generalization of the approach in [Frankot, Chellapa,87] is to generalize the cost function.

The cost function that was minimized was:

$$C = \int (I-R)^2+\lambda(U_{rr}^2+2U_{ry}^2+U_{yy}^2)drdy \qquad (3)$$

where
I = Actual image
R = Reflectance map
ur = Slope in the range direction
uy = Slope in the azimuth direction
urr = Second partial derivative of u in the range direction
ury = Partial derivative of u with respect to r,y directions
uyy = Second partial derivative of u in the y direction

λ= Regularization parameter that controls the amount of curvature in
the derived height model.

A generalization of the algorithm to N images minimizes the following cost function (superscripts refer to image number):

(4)

$$C = \int (I^1-R^1)^2+(I^2-R^2)^2+\ldots (I^n-R^n)^2+\lambda(U_{rr}^2+2U_{ry}^2+U_{y}^2$$

Using the generalized cost function for N images we have [Thomas et al,89]:

$$U_r = \overline{U_r}+\frac{3\varepsilon^2}{10\lambda}((I^1-R^1)R_{u_r}^1+\ldots+(I^n-R^n)R_{u_r}^n) \qquad (5)$$

$$U_y = \overline{U_y}+\frac{3\varepsilon^2}{10\lambda}((I^1-R^1)R_{u_y}^1+\ldots+(I^n-R^n)R_{u_y}^n)$$

In the generalization presented above, it is assumed that the reconstructed terrain model and all images are in the same co-ordinate system. In general, this assumption is not valid, unless the images are rectified and registered. Because of registration problems, this algorithm is not as effective as the one discussed in the next section section.

This model containing the surface facet slopes for each terrain element in the estimated model is converted to a model containing height values using the Fourier projection technique of [Frankot, Chellapa, 87]. This process is performed on every iteration of the complete algorithm. This Fourier- based technique is generalized in the next section.

2.2.2 Variation B: Combining Models Using Fourier Projection

Another approach for generalizing the reconstruction process to N images is to combine single image-derived models in the Fourier projection stage of the algorithm in [Frankot, Chellapa, 87]. This approach is superior to the previous cost function generalization and is discussed below.

In order to handle the case where the different images are in different coordinate systems (e.g. different slant-range projections), an integrability projection based algorithm has been developed. In this approach the slope information (as opposed to pixel values) for each pixel from the different images are combined in the integrability projection stage to derive a composite elevation model. The differences in the co- ordinate systems from one image to the next are handled by a rotational co-ordinate transformation and resampling. The rotational transformation that would be used is exactly the transformation used in [Frankot, Chellapa,87] to approximate the SAR imaging geometry. The following equations are developed for the two image case and are readily extended to the N image case.

In this approach the iterative form for a single image is used to derive a set of slopes for each of the images separately. These slopes are then simultaneously projected on to an integrable surface as follows:

The estimated non-integrable slopes for image (1) are Z_x^1, Z_y^1, and the estimated, non-integrable slopes for image (2) are Z_x^2, Z_y^2.

Let the integrable slopes be Z_x, Z_y.

The projection at each iterate is attained by minizing the following distance function:

(6)

$$d = \int (Z_x-\widehat{Z_x})^2+(Z_x-\widehat{Z_x^1})^2 +(Z_x-\widehat{Z_x^2})^2+(Z_y-\widehat{Z_y^2})^2+(Z_y-\widehat{Z}$$

to get the following equation as shown in [Thomas et al. 89].

$$C=\frac{(C1+C2)P_X+(C2+C4)P_Y}{2(P_X+P_Y)} \quad (7)$$

where

C, C1, C2, C3, C4 represent the Fourier coefficients of Z, Z_x^1, Z_y^1, Z_x^2, Z_y^2 and Z represents the desired integrable surface.

Our experience on synthetic test imagery indicates that convergence was attained with far fewer iterations using the multiple image algorithms than with the single image version.

2.2.3 Reinforcement of Spot Heights

Another modification of the single-image algorithm is the correction of estimated heights in the Fourier projection step described above, to take into account a priori knowledge of some spot heights gained from stereo. Such spot height values represent constraints on the estimated values. Conceivably, such constraints could be reinforced at each iteration.
However, reinforcement at each iteration of a "hard" constraint, such as these spot heights, tends to cause an oscillatory "ringing" in the estimated surface. Therefore, such corrections are presently made in a
"soft" mode. This means that corrections are made using a given fraction of the difference between the constraint value and the estimated value at each iteration. As the iteration number increases, the fraction of the difference between the constraint value and the estimated value can be allowed to increase.

The use of a soft version of a constraint rather than a "hard" constraint is generally useful in iterative optimization problems, since there may be no connected path from a suboptimal solution to the optimal solution which satisfies the hard constraint. However, by continuity, a path exists consisting of suboptimal solutions which satisfy soft constraints.

There were some robustness problems observed in this spot height reinforcement process during the simulations. Near peaks and troughs of a smoothly varying surface, there was an observed "undershoot" and "overshoot" of the constructed surface. The reconstruction process appears to be somewhat ill-conditioned in those regions, since there is almost no difference between predicted and sensed gray-values. The reinforcement, on any iteration, of spot heights in such regions appears to be reduced by the algorithm on subsequent iterations.

Future follow-on efforts will investigate other approaches for reinforcing spot heights within the framework of this algorithm. One possible approach consists of reinforcing entire subregions around the spot height.

2.3 Approach #2: Simultaneous Equation Representation

An entirely different approach to terrain reconstruction involves posing the problem as one of estimating an integrable solution directly, rather than projecting estimated solutions onto subspaces of integrable functions as was done in the previous algorithms.

This method relies on formulating the reconstruction problem as a problem of estimating the coefficients of integrable basis functions. In this case, the basis functions are chosen to be 2-D Fourier Series basis functions.

Using these basis functions, any existing height constraints can be formulated as linear equations for the basis function coeeficients. The predicted image information is formulated as a set of non-linear equations involving the derivatives of the terrain.

An advantage over the previous procedure is that there is a more natural representation of both slope and height information, rather than making height corrections to an algorithm dominated mostly by slope information.
This approach also has the advantage that image registration is not required because it is the terrain Fourier coefficients that are being estimated, not individual terrain cells.

A disadvantage is that a system of nonlinear equations must be solved. One approach is to linearize these equations and use the matrix generalized inverse to minimize the length of the residual vector. The use of the Householder transformation will allow more efficient and stable computation of this inverse.

The cost function in this case is the least-squares norm of the residual vector. The regularization is applied via directly band-limiting the spatial frequencies of the terrain rather than the use of penalty functions.

This procedure has not been implemented, and is still in the conceptual stage.

3. RESULTS

The shape-from-shading algorithm (approach#1 version a) described above was implemented and a number of experiments were conducted to evaluate the performance. The experiments were initially conducted with simulated radar images and later with actual radar images.

Figure 1 is a perspective view of a height model of the Brazzeau range in Canada. The variation in height from the lowest point to the highest is approximately 2000m, the spacing of sample points was 60m, and the size of the height model was 128x128 elements. Using this height model as input, a set of three noise-free radar images were simulated with the radar assumed to be flying at a height of 1000km and imaging with incidence angles of 65.38, 56.44, and 50.28 degrees to the near edge of the scene. Figure 2 shows one of these simulated images.

A three image version of the same shape-from-shading algorithm was run using the three radar images simulated above. Using a heavily smoothed version of the original height model as a starting condition, and by locally enforcing 225 regularily spaced height points, the algorithm was able to generate a height model. On comparing the drieved height model with the actual height model (shown in figure 1), it was found that the standard deviation of the error was 80m. This is shown in figure 3.

The algorithm was then run using a single image as input and using the same initial starting condition (i.e. the same smoothed version of the height model). Figure 4 shows a perspective view of the height model thus derived, and the standard deviation of error on comparing with the true height was found to be 119m.

In order to test the performance of the algorithm with real radar images, a two image version (since only two overlapping images were available at the time) of the same algorithm was implemented. In this experiment, stereo radar images were used to create a height model, and the height model was then used to rectify the images. The rectified and co-registered images were averaged from 6m pixels to 24m pixels. Figures 5 and 6 show the 128x128 pixel input images. The algorithm was also provided the stereo derived height model as a starting model. Figure 7 shows the stereo derived height model as a radar power map (excluding the geometric distortions) created from the same position the true radar created one of the input images (from the left of the image and imaging to the right). Figure 8 shows the output height model derived from the shape-from-shading algorithm presented as a radar power map. Spot-heights were not enforced in this experiment, and due to the lack of map data we were unable to compute a quantitative error.

4. DISCUSSION

As seen in figure 8, the multiple image shape-from-shading is capable of providing a height model that appears to be an improvement from the results provided by radar stereo. Although a map derived height model was unavailable for comparison, the final height model seems to be consistant with the input images. There are however, a few obvious errors in the height model, and these errors manifest themselves in figure 8 as bright points on the side facing the radar and as horizontal dark lines (shadow) on the opposite side. The algorithms described in this paper do not resolve the ambiguity between slope and σ_0, but only finds the most consistent set of slopes that could have resulted in the input images. As a result slopes could be adjusted due to a variation in σ_0 as opposed to a change in the slope, resulting in false heights.

Another artifact that was observed was the ringing effect along the borders of the derived model and it is anticipated that using appropriate window functions during the processing will minimize effect.

5. CONCLUSIONS AND FUTURE PROSPECTS

The above results demonstrate the advantages of using multiple SAR images for shape from shading to supplement ordinary radargrammetry. The possibility exists for combining these methods with other multiple image methods, such as phtometric stereo [Horn,80]. The latter method independently solves the simultaneous set of shading equations, for each registered pixel in a multiple image data set.

During the Magellan Mission to Venus, the SAR images are expected to be of low resolution and hence the parallax induced distortions will be minimal or nonexistent in the overlapping areas. Due to the expected lack of any significant parallax information in these images, stereo techniques will not provide an accurate height model and the overlapping images can be easily registered.

For such a scenario, shape from shading provides a hope for obtaining terrain information. Furthermore, the use of multiple images should decrease the sensitivity to noise due because of the additional information that is available.

REFERENCES

1.Frankot,R., Chellapa,R., "Application of a Shape from Shading Technique to Synthetic Aperture Radar," Proc. Int'l Geoscience and Remote Sensing Symposium, Ann Arbor, MI, May 1987.

2.Horn, B., Robot Vision, MIT Press, 1986.

3.Wildey, R., "Radarclinometry for the Venus Radar Mapper, "Photogrammetric Engineering and Remote Sensing,, Vol. 52, #1, Jan
1986.

4.Thomas,J., et al, "Mutiple SAR Image Shape From Shading", in progress.

5.Baltes, H.P.ed., Inverse Scattering Problems in Optics, Springer, 1980

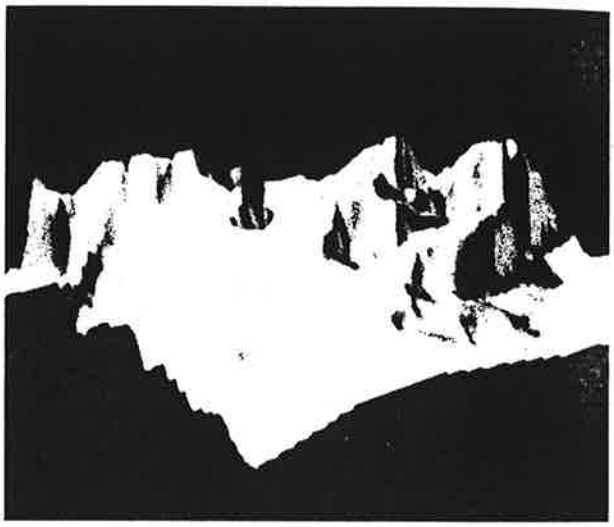

Fig. 1

Fig. 2

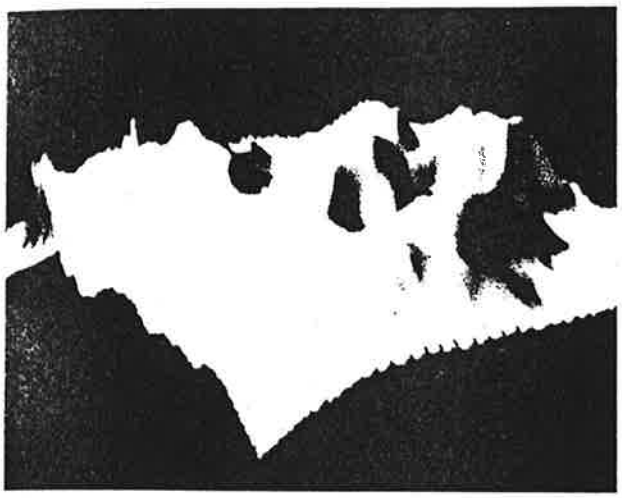

Fig. 3

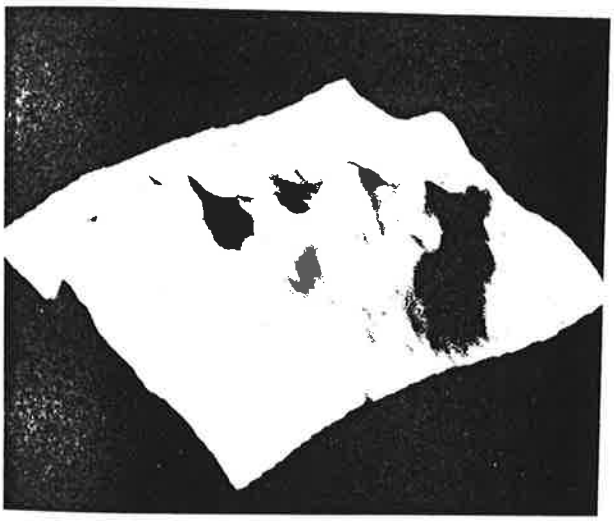

Fig. 4

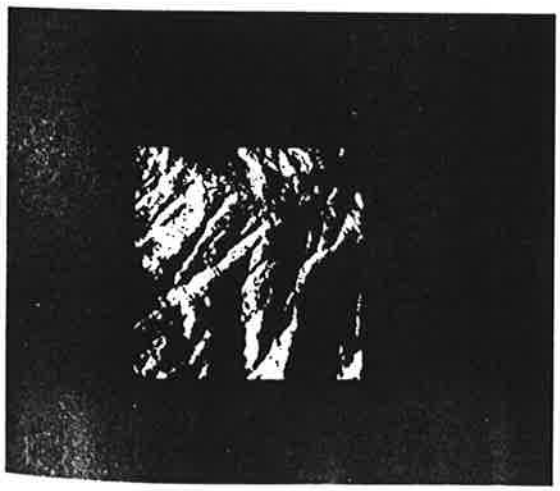

Fig. 5

Fig. 6

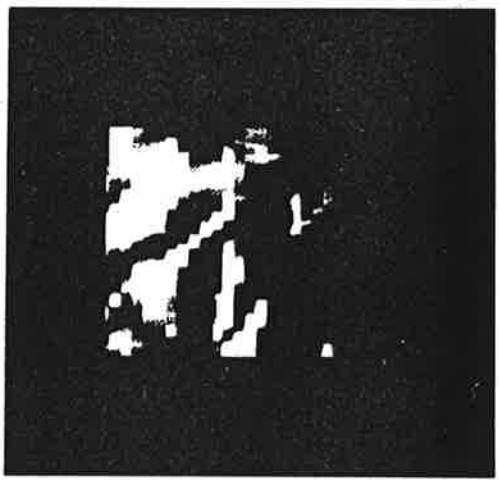

Fig. 7

Fig. 8