DIGITAL SURFACE MODELLING FROM SPOT 5 HRS IMAGERY USING THE AFFINE PROJECTIVE MODEL

C. S. Fraser [a], P.M. Dare [b], T. Yamakawa [a]

[a] Department of Geomatics, University of Melbourne, Victoria 3010, AUSTRALIA
c.fraser@unimelb.edu.au; yamakawa@sunrise.sli.unimelb.edu.au

[b] Airborne Research Australia, Flinders University, PO Box 335, Salisbury South, SA 5016, AUSTRALIA
Paul.Dare@AirborneResearch.com.au

Commission I, WG I/2

ABSTRACT:

This paper presents the results of work carried out under the auspices of the joint ISPRS CNES Scientific Assessment Program of DEM generation from SPOT 5 high resolution stereoscopic (HRS) image data. A stereo pair of HRS images of Melbourne, Australia, with a ground sample distance of 5m in the along track direction, and 10m in the across track direction, was used to create digital surface models of varying grid spacings of the urban area surrounding the city of Melbourne. Conjugate points from the SPOT stereopair were located using intensity based image matching, and triangulated using the affine projective model. The resultant digital surface models were compared to precisely located ground control points and a reference digital terrain model. The results showed that high quality digital surface models can be produced from HRS image data using the algorithms described. The comparison between the surface model and ground control points revealed an RMS height error of 3-5m or 0.6-1 pixel.

1. INTRODUCTION

The purpose of the SPOT 5 Scientific Assessment Program (SAP) is to evaluate the utility of SPOT 5 high resolution stereoscopic (HRS) imagery for generating digital terrain models (DTMs) (Baudoin et al., 2004). Under this program, tests have been carried out by many different research teams on SPOT 5 HRS data for many different regions of the world. The study described in this paper has made use of data acquired over the city of Melbourne, Australia. Since the research teams involved in the SAP are independent, a range of different processing techniques have been brought to bear on the issue of extracting digital surface models (DSMs) from SPOT 5 HRS imagery.

Previous research carried out at the University of Melbourne on the subject of geometric modelling of high-resolution spaceborne sensors has focused on the use of 'alternative' mathematical models, such as the rational polynomial coefficient (RPC) model (Fraser et al., 2002; Hanley et al., 2002) and the affine projective model (Yamakawa et al., 2002, Fraser and Yamakawa, 2004). The success of these studies has led to these models being incorporated into the digital surface modelling algorithms created specifically for high-resolution satellite imaging sensors, such as IKONOS and QuickBird. Results of studies utilising stereopairs of IKONOS images have shown that both the RPC model and affine projective model can be used to create high quality DSMs (Dare, 2004; Dare and Fraser, 2004).

Since SPOT 5 HRS data is not supplied with RPCs, the research presented in this paper concentrates on the use of the affine projective model. Note that traditional photogrammetric techniques can also be used to generate surface models from SPOT 5 data, but such investigations have been left to other members of the SAP.

2. SPOT 5 DATA

The data used in this study covered an area of roughly 140km by 50km around the city of Melbourne, Australia (Figure 1). The image was acquired at 10:20am local time on February 19th, 2003. The land cover types in the scene include water, urban areas, agricultural land and forests. The elevation varies from sea level to around 600m above mean sea level. A portion of the western side of the region was partially obscured by cloud cover during image acquisition, but this made little difference to the generation of the surface models in this study, since these were centred on other areas within the scene.

The pixel size of HRS imagery is 5m in the along track direction, and 10m across track. Since the stereo images were acquired in the same orbit in a fore and aft configuration, the parallax differences occur in the along track direction. Therefore, the along-track pixel size (5m) is relevant when discussing the accuracy of surface models with respect to the pixel size of the original images.

3. AFFINE PROJECTIVE MODEL

Since the launch of the first high-resolution satellites, and even before, much attention has been focused on the use of alternative geometric models, especially RPCs, for restitution, orthorectification and terrain modelling (Dowman and Dolloff, 2000; Dial, 2000; Grodecki, 2001; Grodecki and Dial, 2001). Significant research has also been carried out on other models, such as the direct linear transform (DLT) model and the affine projective (AP) model. An inter-comparison of these three models (RPC, DLT and AP) with IKONOS data showed that although the degree of complexity of the models differs widely, differences in the results can be expected to be small (Fraser et al., 2002, Hanley et al., 2002).
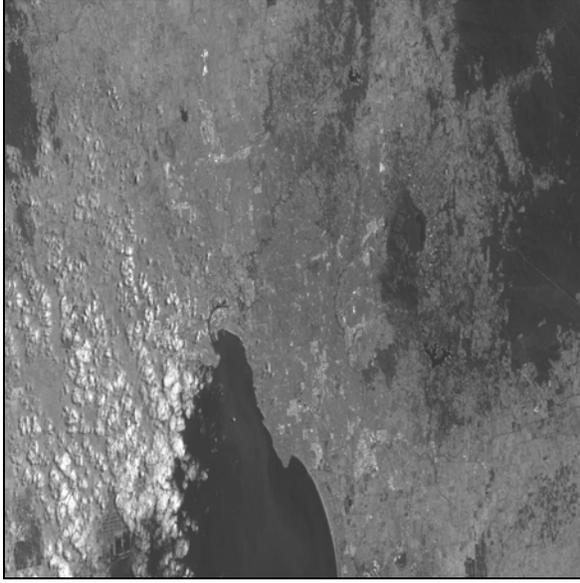
Figure 1. SPOT 5 scene of Melbourne, Australia

The affine projective model is the simplest method of relating image space coordinates to object space coordinates without any knowledge of the sensor model or the exterior orientation of the sensor. Early research was carried out using moderate resolution satellite sensors such as SPOT and MOMS (Okamoto et al., 1999; Hattori et al., 2000), but more recently it has been successfully applied to high resolution satellite imagery, specifically IKONOS (Yamakawa et al., 2002, Fraser and Yamakawa, 2004, Hanley et al., 2002). Although it requires only a modest number of ground control points (GCPs), the affine model has been shown to produce results to sub-pixel accuracy. The general form of the model describing an affine transformation from 3D object space (X,Y,Z) to 2D image space (x, y) for a given point i is expressed as:

$$
\begin{aligned}
x_i &= A_1 X_i + A_2 Y_i + A_3 Z_i + A_4 \\
y_i &= A_5 X_i + A_6 Y_i + A_7 Z_i + A_8
\end{aligned}
\tag{1}
$$

where x, y are image space coordinates; X, Y, Z are object space coordinates; and A_1 to A_8 are the eight affine parameters.

These eight parameters per image account for translation, rotation, and non-uniform scaling and skew distortion. Implicit in Equation 1 are two projections, one scaled-orthogonal and the other skew-parallel. In the reported implementation of the affine projective model, all model parameters are recovered simultaneously along with triangulated ground point coordinates in a process analogous to photogrammetric bundle adjustment.

The affine model assumes, firstly, that the projection from object space to image space is an affine projection and, secondly, that lines of acquired image data are parallel to each other. The first assumption holds true for high-resolution satellite imaging sensors which have a very narrow field of view of around 2° or less. Previous studies have shown the assumption of parallel rather than perspective projection to be

sufficiently valid. The second assumption is true if the satellite travels in a straight line, parallel to the ground during image acquisition. Thus, the Universal Transverse Mercator projection (UTM zone 55) was employed as the object space reference coordinate system in the reported investigation, since the assumption of a straight line track for the satellite, parallel to the 'XY plane', is sufficiently valid within this projection coordinate framework.

4. IMAGE MATCHING

The matching methodology implemented in this study combines image space matching with an object space geometric constraint, namely the affine projective model. Usually image space matching uses a geometric constraint in image space, such as epipolar geometry, to constrain the matching process. The constraint is necessary to reduce the search space, which in turn reduces processing time, as well as reducing the likelihood of erroneous matches. The use of an object space geometric constraint replaces the need for the epipolar constraint.

Matching points using geometric constraints in object space rather than image space is simply another way of describing the search for an unknown height value by moving along an image nadir line until a highly correlated match of image pixels is found. This method of matching has previously been described by Benard (1984), and subsequently incorporated into many object space matching processes (Helava, 1988; Ebner and Heipke, 1988; Gruen and Baltsavias, 1986). The method works by taking an object space point (X_0, Y_0, Z_0) and projecting it, using the affine projective model, into the image spaces of the images being matched: (x_1, y_1) for image 1 and (x_2, y_2) for image 2. These two image points are then matched, in image space, using a typical intensity-based matching strategy. The similarity measure (in this case the cross-correlation coefficient) for the match is recorded. A new object space point (X_0, Y_0, Z_0+dZ) is then transformed into image coordinates and matched as before. Once again the similarity measure is recorded. The process is repeated for all values of Z_i between the lower and upper limits of Z. The value of Z_i which corresponds to the greatest similarity measure is taken as the determined height at the point (X_0, Y_0). The process is repeated for all (X_i, Y_i).

The similarity measure implemented in this matching strategy to compare conjugate points was the cross-correlation coefficient, γ, given by (Gonzalez and Woods, 1992):

$$
\gamma(x,y) = \frac{\sigma_{MS}}{\sigma_M \sigma_S}
\tag{2}
$$

where σ_M and σ_S are the standard deviations of the master and slave chips being matched, and σ_{MS} is the covariance of the intersection of the master chip with the slave chip.

Since the matching strategy in this study is driven by an object space geometric constraint, points have to be initially selected in object space before being transformed into image space and matched. Therefore, in order to generate the candidate matching points, a grid of three dimensional object space points covering the area of interest was created. These points were then sequentially transformed into image space coordinates and matched according to the method described above.

5. RESULTS

Two tests were carried out on an area of 10km by 10km centred on the Melbourne central business district. The height difference throughout the test area was only about 60m (except for large office buildings), with the terrain being as indicated by the DTM shown in Figure 2.
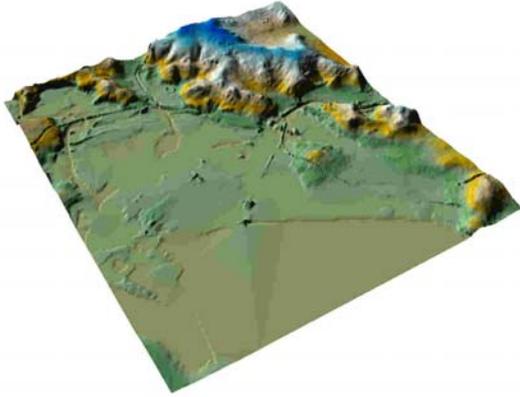

Figure 2. DTM of 10 x 10km Melbourne test area.

The 100km^2 test area was selected since it coincided with a reference DTM and 80 precisely GPS-surveyed GCPs. The difference between the two tests was in the number of GCPs used to calculate the parameters of the affine projective model. Although only four points are required to determine the parameters, a larger number gives a degree of redundancy. Therefore in the first test 10 GCPs were used, and in the second 20 were employed.

Table 1 shows the number and percentage of successful matches for each test. In this study, successful and unsuccessful matches have been differentiated from each other according to the values of the cross-correlation coefficient.

No. of GCPs	Number of matched points	γ > 0.8	γ > 0.9
10	43572	34352 (78.84%)	21119 (48.27%)
20	43894	34597 (78.82%)	21359 (48.66%)

Table 1. Number of successful matches.

It can be seen from Table 1 that there is no significant difference between the proportion of successful matches for each test. This is a very encouraging result since it indicates that reducing the number of GCPs from 20 to 10 does not impact negatively on the results. From a practical point of view this is very important, since collection of high quality ground control can be a very time consuming process. A further point to notice is that the test results endorse the use of the matching strategy used in this study: in both cases nearly 50% of the matched points have a correlation coefficient greater than 0.9.

Results of each of the tests were further analysed by comparing the generated surface model with both a reference terrain model, and ground surveyed check points. These results are presented in Tables 2 and 3 below. There were 60-70 surveyed check points and around 110 DTM check points used.

10 GCPs	γ > 0.8		γ > 0.9	
	RMS (m)	σ (m)	RMS (m)	σ (m)
Difference: DSM vs. check points	4.11	4.14	3.91	3.94
Difference: DSM vs. DTM	5.59	5.62	4.96	4.98

Table 2. Height difference between DSM and reference data for test with 10 GCPs.

20 GCPs	γ > 0.8		γ > 0.9	
	RMS (m)	σ (m)	RMS (m)	σ (m)
Difference: DSM vs. check points	3.18	3.21	2.96	3.01
Difference: DSM vs. DTM	4.78	4.80	3.96	3.98

Table 3. Height difference between DSM and reference data for test with 20 GCPs.

Tables 2 and 3 show the RMS height differences, and the standard deviations (σ) of those differences, between the derived surface models and the reference data, for two different groups of triangulated points: those with cross-correlation coefficient values greater than 0.8, and those with values greater than 0.9. As would be expected, the groups of triangulated points with higher cross-correlation coefficients give a better surface representation than those with lower cross-correlation values. Even so, the differences between the groups are quite small, with this result being consistent for both tests.

The points used for the comparison between the surface model and the reference data were chosen carefully so that the height differences could be measured in regions unaffected by ground features such as buildings and vegetation cover. The fact that the differences between the surface models and the check points are less than the differences between the surface models and the DTM is difficult to explain, but is most likely due to errors in the reference DTM (either relative or absolute), which was created from stereo aerial photography and required significant manual editing.

Since the GPS surveyed check points are obviously the most reliable reference, they have provided the best yardstick against which to assess the SPOT 5 surface model. With just 10 GCPs to calculate the parameters of the affine projective model, the RMS differences were of the order of 4m, which equates to 0.8 pixels. When 20 GCPs were used to calculate the affine projective parameters, the RMS height differences were around 3m, or 0.6 pixels.

6. CONCLUDING REMARKS

This study, which has been conducted under the auspices of the SPOT 5 HRS-SAP, has examined the accuracy of DSMs created from SPOT 5 HRS imagery via an affine projective sensor orientation model. The affine model has previously been shown to produce highly successful geopositioning and DSM results with other high-resolution satellite imaging systems.

The results obtained have shown that, at least for a 10km by 10km area in the centre of a pair of SPOT 5 HRS images, DSMs with a vertical accuracy of 3-5m, or about 0.6-1 pixel, can be obtained. No information on the sensor model or satellite ephemeris is required to attain this accuracy, but a modest number of well-measured GCPs are required to determine the parameters of the affine model. In this case, configurations of 10 and 20 GCPs were considered. The results obtained indicate very little accuracy distinction in the final DSMs for the cases of 10 and 20 GCPs.

The scope of the work presented has been limited in that no tests have thus far been carried out with full-scene images. Future work will investigate the performance of the affine model over full-scene HRS images, where the effective field of view is substantially larger than the 1^0-2^0 range in which the assumptions implicit in this linear model have been shown to hold up well for all practical purposes.

7. ACKNOWLEDGEMENTS

8. REFERENCES

Baudoin, A., Schroeder, M., Valorge, C., Bernard, M. and Rudowski, V., 2004. The HRS-SAP initiative: A scientific assessment of the High Resolution Stereoscopic instrument on board of SPOT 5 by ISPRS investigators. *International Archives of Photogrammetry and Remote Sensing*, 34(1), (this volume), 7 pages.

Benard, M., 1984. Automatic stereophotogrammetry: a method based on feature detection and dynamic programming. *Photogrammetria*, 39(4-6): 169-181.

Dare, P. M., 2004. Investigation of geometric constraints for matching high resolution satellite images. *International Archives of Photogrammetry and Remote Sensing*, 34(3), (this volume).

Dare, P. M. and Fraser, C. S., 2004. Digital surface modelling from commercial high resolution satellite imagery. *Geomatics Research Australasia* (in Press).

Dial, G., 2000. Ikonos satellite mapping accuracy. *Proc. ASPRS Annual Conference*, Washington D.C., 22-26 May, 8 pages (on CD-ROM).

Dowman, I. J. and Dolloff, J. T., 2000. An evaluation of rational functions for photogrammetric restitution. *International Archives of Photogrammetry and Remote Sensing*, 33(3): 254-266.

Ebner, H and Heipke, C., 1988. Integration of digital image matching and object surface reconstruction. *International Archives of Photogrammetry and Remote Sensing*, 27(2), pp. 534-545.

Fraser, C.S. & T. Yamakawa, 2004. Insights into the affine model for satellite sensor orientation. *ISPRS Journal of Photogrammetry and Remote Sensing* (in Press).

Fraser, C. S., Hanley, H. B. and Yamakawa, T., 2002. Three-dimensional geopositioning accuracy of IKONOS imagery. *Photogrammetric Record*, 17(99): 465-479.

Gonzalez, R. C. and Woods, R. E. (1992). *Digital image processing*. Addison-Wesley, New York, 716 p.

Grodecki, J., 2001. IKONOS stereo feature extraction - RPC approach. *Proc. ASPRS Annual Conference,* St. Louis, 23-27 April, 7 pages (on CD-ROM).

Grodecki, J. and Dial, G., 2001. IKONOS geometric accuracy. *Proc. Joint ISPRS Workshop "High Resolution Mapping from Space 2001",* Hanover, 19-21 Sept., 10 pages (on CD-ROM).

Gruen, A. and Baltsavias, E., 1986. High precision image matching for digital terrain model generation. *International Archives of Photogrammetry and Remote Sensing*, 26(3), pp.284-296.

Hanley, H.B., Yamakawa, T. & C.S. Fraser, 2002. Sensor orientation for high-resolution satellite imagery. *International Archives of Photogrammetry and Remote Sensing,* 34(1): 69-75 (on CD-ROM).

Hattori, S., Fraser C. S., Ono, T. and Hasegawa, H., 2000. An alternative approach to the triangulation of SPOT imagery. *International Archives of Photogrammetry and Remote Sensing*, 32(4), pp. 359-366.

Helava, U. V., 1988. Object space least squares correlation. *Photogrammetric Engineering and Remote Sensing*, 54(6), pp. 711-714.

Okamoto, A., Ono, T., Akamatsu, S., Fraser, C.S., Hattori, S. and Hasegawa, H., 1999. Geometric characteristics of alternative triangulation models for satellite imagery. *In Proceedings of the ASPRS Annual Conference*, Portland, Oregon, 17-21 May 1999.

Yamakawa, T., Fraser, C.S. and H.B. Hanley, 2002. High-precision 3D ground point positioning from IKONOS satellite imagery. *Journal of Japan Society of Photogrammetry & Remote Sensing,* 41(2): 36-43.