DIGITAL ORTHOMOSAICS AS A SOURCE OF CONTROL FOR GEOMETRICALLY CORRECTING HIGH RESOLUTION SATELLITE IMAGERY

Paul Dare*
Nicki Pendlebury**
Clive Fraser***

*Department of Geomatics, University of Melbourne, Victoria 3010, AUSTRALIA
Telephone: +613 8344 6806; Fax: +613 9347 2916
Email: pmdare@unimelb.edu.au

**Department of Geomatics, University of Melbourne, Victoria 3010, AUSTRALIA
Telephone: +613 8344 6806; Fax: +613 9347 2916
Email: nicpendlebury@optusnet.com.au

***Department of Geomatics, University of Melbourne, Victoria 3010, AUSTRALIA
Telephone: +613 8344 6806; Fax: +613 9347 2916
Email: c.fraser@unimelb.edu.au

KEYWORDS: High resolution satellite imagery; orthomosaics; image matching

ABSTRACT: Recent research conducted at the University of Melbourne has shown that commercial high resolution satellite imagery can be accurately rectified using the combination of bias-corrected rational polynomial coefficients (RPCs) and ground control points (GCPs). Research carried out to date has used GCPs measured using differential GPS, a relatively expensive and time-consuming process. This paper presents the initial results of an investigation into the possibilities of using GCPs automatically extracted from digital orthomosaics created from aerial photography. As long as common points can be located in corresponding pairs of orthomosaics and satellite images, the need for ground surveyed GCPs is eliminated. Experiments using various combinations of matched points extracted from aerial photography and high resolution Ikonos satellite images were used to remove positioning biases in RPCs and hence improve the geometric precision of the satellite images. The results showed that, in most scenarios, orthomosaics prove to be an excellent source of control for the geometric correction of high resolution satellite imagery.

1. INTRODUCTION

It has been over three years since the successful launch of Ikonos, the first commercial high resolution imaging satellite. Since then other high resolution satellites, such as QuickBird and EROS have also been launched. During this time a significant amount of research has been carried out into the geometric processing of high resolution satellite sensors. Even though the camera model is not available for the Ikonos sensor, research has shown that one of the options for accurate three dimensional measurement is provided by bias-corrected rational polynomial coefficients (RPCs) using as few as one or two ground control points (GCPs) (Fraser et al., 2002). The aim of the study described in this paper is to assess the potential of the combination of an orthomosaic and a digital terrain model (DTM) as a source of control.

The difference in costs between *Geo* and *Precision* Ikonos image products are significant (Space Imaging, 2002). Nevertheless, with a few precisely measured GCPs, accurate three dimensional measurements can be made from the base level Ikonos data product. However, the accuracy of the results is dependent on the precision of the GCPs. Although differential GPS provides a direct method of accurate GCP measurement, costs can be high if a large number of points are required. Furthermore, the differential GPS measurement equipment must actually be available in the region where the GCPs are to be surveyed. An alternative to the use of differential GPS measurements is to use control from another image source, ideally an orthomosaic. The advantages that this approach offers are firstly that control points can be found automatically or semi-automatically using image matching techniques, and secondly that costs are fixed, irrespective of the number of control points found.

This paper addresses some of the issues associated with using an orthomosaic as a source of control. In doing so it also explores the challenge of multi-source, multi-resolution data fusion: in order to achieve the goal of automatic multi-sensor data integration, the problems of tie point detection must be resolved. The automatic point matching presented in this paper represents a first step in the fusion of very high resolution spatial data sets.

2. DATA USED IN THIS STUDY

The data used in this study included high resolution Ikonos *Geo* satellite imagery (both nadir and stereo images), very high resolution orthomosaics and an accurate DTM. The test site for the study was a region a couple of kilometres to the north of the Central Business District of Melbourne, Australia. The high resolution satellite imagery was acquired from the Ikonos satellite in two configurations, nadir and stereo (fore and aft), and supplied at a ground sample distance (GSD) of 1m.

Previous studies have shown that the internal accuracy of Ikonos images is very good, being less than half a metre in both easting and northing for all images, even though there is a height variation over the Melbourne test site of some 50m (Hanley and Fraser, 2001). Absolute accuracies were initially quite poor due to the fact that the Ikonos imagery was supplied as a *Geo* product. This product type has the lowest accuracy specification (RMS positional accuracy of 25m), but is the cheapest to purchase.

The high resolution colour orthomosaic used in this study was supplied with a pixel size of 0.15m. An assessment into the relative and absolute planimetric accuracy was carried out as an associated study (Dare and Fraser, 2002), the results of which are summarised in table 1. The orthomosaic was found to be close to published accuracy specifications of 0.15m, and hence sufficiently accurate for use as a source of planimetric control for geometrically correcting the Ikonos imagery.

	RMSE in easting (m)	**RMSE in northing (m)**
Absolute accuracy	0.22	0.17
Relative accuracy	0.18	0.17

Table 1. Relative and absolute planimetric accuracy of the orthomosaic (after Dare and Fraser, 2002)

The DTM used as the source of vertical control in this study had a grid spacing of 25m. To determine its absolute accuracy, an assessment was performed where heights interpolated from the DTM were compared with GPS measured ground control. It was found that the RMS discrepancy was 0.24m, with a range of –0.16m to 0.46m. Consequently it was decided that this DTM would provide sufficiently accurate height information for this study.

3. TWO DIMENSIONAL GEOMETRIC CORRECTION

Hanley and Fraser (2001) performed a two dimensional geometric correction of the Ikonos data using 41 ground surveyed GPS points, the results of which are shown in table 2. Following on from this work, ground control measured from the orthomosaic was used to perform the same two dimensional geometric correction of the Ikonos imagery. Using a total of 27 tie points and least squares estimation, the parameters of two transformation functions, similarity and affine, were calculated. The residuals of the transformed points are also shown in table 2.

Source of control	Image	Similarity		Affine	
		X (m)	Y (m)	X (m)	Y (m)
Ground surveyed GPS	Ikonos nadir	0.27	0.33	0.26	0.32
	Ikonos stereo (left)	0.35	0.39	0.33	0.37
	Ikonos stereo (right)	0.39	0.37	0.38	0.36
Orthomosaic and DTM	Ikonos nadir	0.30	0.31	0.28	0.26
	Ikonos stereo (left)	0.52	0.54	0.48	0.51
	Ikonos stereo (right)	0.44	0.30	0.37	0.25

Table 2. RMS residuals of transformed points.

Comparison of values in table 2 shows that, in two dimensions, the orthomosaic has proved to be an equally good source of control as the GPS measured points in all three images. Hence, in two dimensions, the use of the orthomosaic as a source of control is justified.

4. THREE DIMENSIONAL GEOMETRIC CORRECTION

Since the interior and exterior orientation parameters of the Ikonos sensor at the time of image acquisition are not available, it is not possible to apply traditional photogrammetric techniques to Ikonos image data. However, with the provision of rational polynomial coefficients (RPCs) is it possible to recover three dimensional object space measurements from overlapping stereo imagery (Dowman and Dolloff, 2000). RPCs supplied with Ikonos *Geo* image products give limited absolute accuracy, as would be expected from the specification of the *Geo* imagery. Inaccuracies are manifested in the form of shifts, or biases, in *X*, *Y* and *Z*. The addition of a single GPS-measured ground control point to shift the model in *X*, *Y* and *Z* and remove the biases has been shown to improve the results dramatically (Fraser et al., 2002). Absolute accuracies of approximately 0.5m in *X* and *Y*, and 0.5–1.0m in *Z* have been achieved.

In this study, as with the two dimensional processing described above, the three dimensional geometric correction of the Ikonos imagery was carried out using the orthomosaic as the source of planimetric control, and the DTM as the source of the vertical control. A total of 12 points in the nadir and stereo images were measured. Although only a single point is required to compute the biases and perform the shift, the use of extra points gives added redundancy and hence extra confidence in the result. The RMS discrepancies of check points calculated using the bias-corrected RPCs are given in table 3, as are the corresponding values determined using GPS surveyed ground control.

Source of control	Number of GCPs	Number of check points	RMS discrepancies		
			X (m)	Y (m)	Z (m)
Ground surveyed GPS	4	36	0.41–0.67	0.44–0.69	0.58–0.90
Orthomosaic and DTM	4	8	0.21–0.29	0.29–0.41	0.56–0.79

Table 3. RMS discrepancies of check points calculated using bias-corrected RPCs.

Table 3 shows that in the case of three dimensional point measurement, ground control determined from the orthomosaic and the DTM gives a solution comparable to that which can be achieved using GPS surveyed ground control. In each case, the same number of GCPs have been used to calculate the RPC bias corrections, but due to the difference in the total number of points measured, fewer check points have been used over a much smaller area to determine the RMS discrepancies corresponding to the orthomosaic GCPs. This would explain why these discrepancies are marginally smaller than those corresponding to the GPS measured GCPs.

5. AUTOMATIC MATCHING

The initial aim of this project was to show that measurements of ground control taken from orthophotography can not only be used to geometrically correct high resolution satellite data, but that the accuracy of the results are at least as good as, if not better than, what can be achieved with high accuracy differential GPS. The results presented above in sections 3 and 4 have shown that this is indeed the case. The ultimate goal of the project however is to exploit the fact that image points can be automatically matched, thus reducing the level of human intervention in the selection of control points. The remainder of this paper describes an automatic image matching technique employed to locate conjugate points in the Ikonos imagery and the orthomosaics to a greater precision than can be achieved by manual measurement. Implementation of a semi-automatic strategy is shown to provide the optimum balance between precision and processing time.

5.1 Matching strategy

There are two steps in the point matching procedure employed in this study: the extraction of potential point features to be matched, and the matching of those point features to give accurate conjugate pairs. In fully automatic matching, both of these steps would be carried out automatically. However, for the initial experiments carried out in this study, the first step was performed manually and the second was performed automatically. Thus the resulting strategy is one of semi-automation. The advantage of the semi-automatic approach is that, for each pair of points being matched, convergence to the correct matching solution is rapid and likely to be blunder-free.

The matching strategy used in determining the conjugate points was based on intensity-based matching and utilised the cross-correlation coefficient, $\tilde{a}$, as a similarity measure (Gonzalez and Woods, 1992). The definition of $\tilde{a}$ is given by

$$g(x,y) = \frac{\sigma_{MS}}{\sigma_M \sigma_S} \tag{1}$$

where $\acute{o}_M$ and $\acute{o}_S$ are the standard deviations of the master and slave chips being matched, given by

$$\sigma_M = \sqrt{\frac{\sum_i \sum_j \left(g_M(x_i, y_j) - \bar{g}_M\right)^2}{mn-1}} \text{ and } \sigma_S = \sqrt{\frac{\sum_i \sum_j \left(g_S(x_i, y_j) - \bar{g}_S\right)^2}{mn-1}} \tag{2}$$

and $\acute{o}_{MS}$ is the covariance of the intersection of the master chip with the slave chip, given by

$$\sigma_{MS} = \frac{\sum_i \sum_j \left(\left(g_M(x_i, y_j) - \bar{g}_M\right)\left(g_S(x_i, y_j) - \bar{g}_S\right)\right)}{nm-1} \tag{3}$$

This was the most appropriate strategy due to the fact that the images had similar radiometric distributions, and that they were already aligned to the same local coordinate system. Had there been geometric differences between the images, an alternative strategy that could model mild geometric distortions, such as least squares matching, would have been required.

5.2 Point extraction

If point extraction were to be performed automatically, an interest operator, such as the Foerstner operator (Foerstner and Gülch, 1987), or the Moravec operator (Moravec, 1977) would have to be implemented. These operators are able to extract points from pairs of images that are likely to have a high rate of being successfully matched. However, the success of resultant matching, and hence the success of the interest operator, relies on the pair of images having a very similar radiometry. Although the radiometry of the Ikonos imagery and the orthomosaic are similar enough for the cross-correlation coefficient to be used as a similarity measure, they are not similar enough for an interest operator to extract a large number of conjugate point features to be used in the matching. Since the rejection rate in automatic point matching is necessarily high, a large number of potential conjugate point features are required. It is for this reason that manual point extraction was used.

5.3 Point matching

There are three parameters associated with point matching based on the cross-correlation coefficient: the chip size, the search space and the step size. These are explained in figure 1.

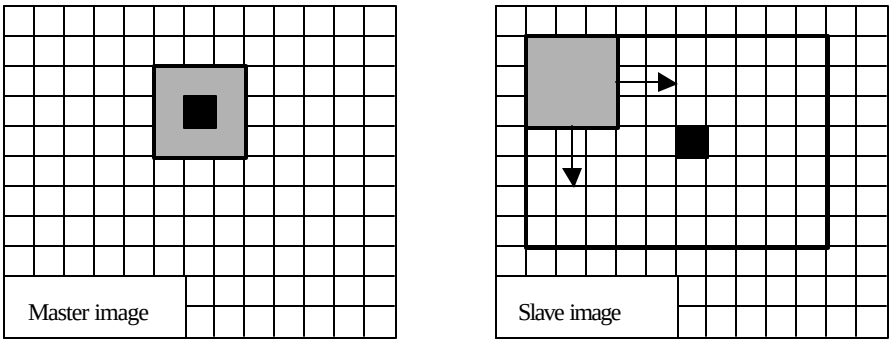

Figure 1. Point matching strategy.

Figure 1 (left) shows a representation of the master image. The solid black square is the pixel to be matched, and the square grey region is the size of the extracted image chip, centred on the matching pixel. In the matching process the image chip from the master image is projected into the slave image. The cross-correlation coefficient is calculated using the pixels where the chip intersects with the slave image. In order to find the best match, the chip is repeatedly projected into the slave image at various locations around the manually selected candidate match point (the solid black square in figure 1 right). The size of the region in which cross-correlation coefficients are determined is known as the search space. As the chip is scanned through the search space in the slave image, the cross-correlation coefficient is determined at discrete locations. The spatial frequency of those locations is the step size. Each time a pair of points is matched, values for these three parameters, the chip size, the search space and the step size, must be specified.

The chip size is the first parameter that needs to be specified. A large chip will have a more distinctive spectral signature than a small chip, greatly increasing the possibilities of a successful (rather than an erroneous) match. Unfortunately with a larger chip, geometric distortions and radiometric discontinuities between the images play a much greater role, leading to the possibility of no match at all. In addition, a large chip size adds significantly to processing time. Therefore a balance between chip size (which must be maximised) and successful matching must be found.

The second parameter to be specified is the search space. This parameter signifies the level of confidence in the initial approximate locations of the candidate matching points. If it is likely that the initial approximate locations are close to the true matching locations, then only a small search space is required. However, if it is possible that the initial approximate locations are far from the true locations, a much larger search space will be necessary. A larger search space implies much greater processing time and a greater possibility of the resultant match being erroneous. A small search space necessarily requires good initial approximate locations of the potential matching points. Since the initial points are found manually, the search space does not need to be very large at all. Conversely, for automatic matching algorithms a large search space is generally required.

The final parameter that needs to be specified is the step size. This value represents the distance between locations at which the cross-correlation coefficient is calculated. It therefore defines the final precision of the coordinates of the matched point. A small step size (high spatial frequency) will give a more precise solution, but at the expense of greater processing time.

Pyramid matching is a technique often used in image matching in order to reduce processing time. Pyramid matching was incorporated into the matching strategy used in this study. In the first layer of the pyramid a wide search space is used with a large step size. In subsequent layers of the pyramid, both the search space and the step size are reduced. Thus in each layer of the pyramid the matching results become more precise, but processing time remains reasonable. The number of layers used in the pyramid matching was generally three to four, depending on the quality of the match.

5.4 Results

The images (the orthomosaic and the Ikonos nadir image) were initially resampled to the same GSD, and the orthomosaic was converted to greyscale, before candidate matching points were located manually in the images. The search space, chip size and step size were then specified.

Initially (i.e., the first layer of the pyramid) the search space was set to 5 pixels in both directions (row and column), and the step size was set to 0.5 pixels. Thus it was necessary to calculate the cross-correlation coefficient at 100 discrete locations, which in terms of computation time, is not unreasonable, so long as the chip size is not too large. In all the tests a chip size of 5 pixels by 5 pixels was used. This process was then repeated using the matching result as the initial approximate matching point. However, this time the search space was reduced, as was the step size; the chip size remained constant. The result of this second iteration (the second layer of the pyramid) was a further refined matching point with a higher cross-correlation coefficient. This process was repeated until the cross-correlation no longer increased, or until enough iterations had been completed (usually a maximum of four).

Once a wide distribution of matched points had been found across the area of overlap between the two images, geometric processing using the RPCs could be performed to assess the quality of the matches. As before, the planimetric coordinates of the points were read from the orthoimage, and the height values were interpolated from the DTM. This time, however, the geometric processing was slightly different since conjugate points had been located in a single Ikonos image (the nadir image) and the orthomosaic. The RPCs associated with the Ikonos image allow the mapping of points from object space to image space, but not *vice versa* (unless the height coordinate is specified). Therefore, bias corrections for the RPCs can be calculated, but image points cannot be projected into object space. The result is that the accuracy assessment is only possible in two dimensions, x and y in image space, as opposed to three dimensions: X, Y, Z in object space. However, since the pixel size is known to be 1m, the results can be expressed in object space coordinates.

In total, 40 matched points were found in the orthomosaic and Ikonos image. The cross-correlation coefficients ranged from 0.563 to 0.996. Of those 40 points, 33 had cross-correlation coefficients of greater than 0.9. Table 4 shows the RMS discrepancies in image space after all the matched points were used to determine the bias corrections to the RPCs. Table 4 also shows the effect of the cross-correlation coefficient on the discrepancies.

Points used	**Number of control points**	**RMS discrepancies**	
		X (m)	Y (m)
All points	40	0.64	0.60
Points with $\gamma > 0.95$	25	0.38	0.56
Points with $\gamma < 0.95$	15	0.92	0.63

Table 4. RMS discrepancies of automatically matched control points

Table 4 shows that using all 40 automatically matched control points gives a RMS discrepancy of 0.64m in X and 0.60m in Y. The magnitudes of the discrepancy values are approximately double those achieved from manually measured control points. However, the group of 40 points includes some matches with cross-correlation coefficients as low as 0.56. If all matched points with poor cross-correlation coefficients are rejected, then the result improves dramatically. It can be seen from table 4 that by only using matched points with cross-correlation coefficients greater than 0.95, the RMS discrepancy in X reduces to 0.38m. In Y there is also a slight improvement. Conversely, by only using comparatively poorly matched points (those with cross-correlation coefficients of less than 0.95), the quality of the result decreases. The overall conclusion is that automatically matched points can yield accuracies to half a metre or better, and that the cross-correlation coefficient gives a good indication of the quality of the match.

6. DISCUSSION

The experiments carried out in this study have shown that the combination of an accurate orthomosaic and a DTM represent a viable source of control in the geometric processing of high resolution satellite imagery. In both two and three dimensional geometric processing, the accuracies of transformed points are similar, irrespective of whether the source of GCPs is the orthomosaic or GPS-measured points. In the automatic matching the results are almost as good, even though a rather basic matching algorithm has been used. In this example it is unlikely that a more advanced automatic matching procedure could provide better accuracies since the precision of the GPS-measured ground control was extremely high. Even so, the results presented in this paper augur well for continued research into automatic geometric processing of high resolution satellite imagery based on image matching. Future work will focus on further automating the procedure of conjugate point detection and matching, in order to support automatic algorithms for data fusion.

7. ACKNOWLEDGEMENTS

The orthomosaic image product used in this study was kindly supplied by Sinclair Knight Merz of Canberra, Australia. Funding was provided by the Australian Research Council.

8. REFERENCES

Dare, P. M. and Fraser, C. S. (2002). An assessment of the geometric quality of AUSIMAGE digital orthophotography. *The Australian Surveyor*. In press – due for publication in 2002.

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.

Foerstner, W. and Gülch, E. (1987). A fast operator for detection and precise location of distinct points, corners and centres of circular features. *ISPRS Intercommission Workshop on Fast Processing of Photogrammetric Data*. Interlaken, Switzerland, pp. 281-305.

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.

Hanley, H. B. and Fraser, C. S. (2001). Geopositioning accuracy of Ikonos imagery: indications from 2D transformations. *Photogrammetric Record*, 17(98): 317-329.

Moravec, H. P. (1977). Towards automatic visual obstacle avoidance. In *Proc. 5th International Joint Conference on Artificial Intelligence*, Cambridge, Massachusetts, U.S.A., pp. 584-592.

Space Imaging (2002). http://www.spaceimaging.com. [Accessed 26th September 2002].