GEOREFERENCING AND ORTHOIMAGE GENERATION FROM LONG STRIPS OF ALOS IMAGERY

Franz Rottensteiner[(1)]**, Thomas Weser**[(2)]**, Clive S. Fraser**[(2)]

[(1)] *Department of Photogrammetry and GeoInformation, Leibniz University Hannover, Nienburger Straße 1, D-30167 Hannover, Germany, Email: rottensteiner@ipi.uni-hannover.de*
[(2)] *Cooperative Research Centre for Spatial Information, Department of Geomatics, University of Melbourne, VIC 3010, Australia, Email: {tweser, c.fraser}@unimelb.edu.au*

ABSTRACT

Precise georeferencing is one of the prerequisites for orthoimage generation from high-resolution satellite imagery. This involves the precise estimation of the parameters of the sensor model that relate the object coordinate system to the row and column indices of the original image and requires the availability of a small number of ground control points (GCPs) that have to be visible in each scene. In this paper, it is shown how the number of GCPs required for the precise georeferencing of ALOS imagery can be reduced by up to 90% using a generic pushbroom sensor model and strip adjustment. Using this technique, the productivity in orthoimage generation can be increased by 300% while still achieving an accuracy of 1 pixel or better.

1. INTRODUCTION

Imagery from the ALOS PRISM and AVNIR-2 sensors offers the potential of orthoimage generation to support medium- and small-scale mapping. One of the prerequisites to exploit the full metric quality of the images is precise georeferencing. In the past we have shown how this can be achieved for single scenes using a generic pushbroom scanner model and a few well-defined ground control points (GCPs) in each scene [1]. GCPs are required to determine corrections for the orbit information delivered with ALOS imagery. In [2] the model was expanded by correction terms for the relative alignment of the individual CCD chips of PRISM level 1B1 images to improve the accuracy of georeferencing by self-calibration.

From the point of view of a mass-production of orthoimages, productivity is limited by the ground control requirements. For instance, Geoscience Australia (GA) distribute orthoimages generated from ALOS PRISM and AVNIR-2 imagery. The traditional production sequence has involved the orientation of every single scene using 10 or more GCPs. The requirement to provide these GCPs in remote areas creates a bottleneck in the production line. This is a very important issue for GA, given that Australia has a land area of 7.6 million km^2 and a single Mode 1 PRISM scene covers only around 1500 km^2.

In order to optimize the economy and productivity of these operations, it is desirable to reduce the number of GCPs required for precise georeferencing. This can be achieved by combining individual ALOS scenes to image strips with shared orientation parameters that are determined simultaneously by strip adjustment. In the strip adjustment, only a few GCPs at the beginning and the end of the strip are needed. This paper describes a computational system that enables realisation of simultaneous georeferencing of ALOS PRISM or AVNIR-2 scenes. It overviews the computational models and steps involved and describes how the workflow has been accomplished within the *Barista* software system [3]. First, the mathematical models used for georeferencing are presented, and the issue of self-calibration is briefly discussed. After that, a fully automated method for GCP determination is described. Finally, the results of an experimental test for strip adjustment are presented, demonstrating how the number of GCPs can be reduced by strip adjustment and also showing the limitations of the automatic technique for GCP definition.

2. SENSOR MODELS

2.1. The Generic Sensor Model

The physical model of the pushbroom satellite imaging process relates a point $\mathbf{P}_{ECS} = (X_{ECS}, Y_{ECS}, Z_{ECS})^T$ in an earth-centred object coordinate system to the position of its projection $\mathbf{p}_I = (x_I, y_I, 0)^T$ in an image file coordinate system. A pushbroom scanner records each image row consecutively at time t while flying over the ground. The coordinate y_I of an observed image point therefore corresponds with the recording time t by $t = t_0 + \Delta t \cdot y_I$, where t_0 is the time of the first recorded image row and Δt the time interval for recording a single image row. The framelet coordinate system refers to an individual CCD array and an image point in framelet coordinates can be expressed as $\mathbf{p}_F = (x_F, y_F, z_F)^T = (x_I, 0, 0)^T$. Each recorded image row is a central projection recorded at time $t(y_I)$. The relation between an observed image point $\mathbf{p}_F$ in and the object point $\mathbf{P}_{ECS}$ is described by Eq. 1:

$$\mathbf{p}_F = \mathbf{c}_F - \delta\mathbf{x} + \lambda \cdot \mathbf{R}_M^T \cdot \{\mathbf{R}_P^T(t) \cdot \mathbf{R}_O^T \cdot [\mathbf{P}_{ECS} - \mathbf{S}(t)] - \mathbf{C}_M\} \quad (1)$$

In Eq. 1, $\mathbf{c}_F = (x_F{}^C, y_F{}^C, f)$ describes the position of the projection centre in the framelet coordinate system; its coordinates are usually referred to as the parameters of interior orientation: the principal point $(x_F{}^C, y_F{}^C)$ and the focal length f. The vector $\delta\mathbf{x}$ formally describes corrections for systematic errors such as velocity aberration and atmospheric refraction. It can also be expanded to model camera distortion or other systematic effects. The shift $\mathbf{C}_M$ and the rotation matrix $\mathbf{R}_M$ describe a rigid motion of the camera with respect to the satellite. They are referred to as the *camera mounting* parameters. Since each image row is recorded consecutively while the satellite is moving, the row also has its own exterior orientation corresponding to the acquisition time t. The satellite orbit path is modelled by time-dependant functions $\mathbf{S}(t) = [X(t), Y(t), Z(t)]^T$. The attitudes of the satellite orbit are described by a concatenation of a time-constant rotation matrix $\mathbf{R}_O$ and a matrix $\mathbf{R}_P(t)$ parameterised by time-dependant functions describing three rotation angles, *roll(t)*, *pitch(t)* and *yaw(t)*. The components of the orbit path and the time-dependant rotation angles are modelled by cubic spline functions. $\mathbf{R}_O$ rotates from the earth-centred coordinate system to a system that is nearly parallel to the satellite orbit path and can be computed from the satellite position and velocity at the scene centre.

2.2. The PRISM Sensor

A speciality of ALOS PRISM is that depending upon the imaging mode, four or six CCD chips are used to record a scene [4]. This results in four or six sub-images that are delivered as separate image files for raw (levels 1A and 1B1) data. These sub-images share their exterior orientation and camera mounting parameters and the focal length. However, each CCD chip has its own framelet coordinate system, and thus the coordinates of the principal point can be different for each of the sub-scenes [2]. In order to model this specific configuration, the bias correction vector $\delta\mathbf{x} = (\delta x, \delta y, 0)^T$ in Eq. 1 is used. We still assume the coordinates of $\mathbf{c}_F = (x_F{}^C, y_F{}^C, f)$ to be identical for all sub-scenes. The corrections $(\delta x_i, \delta y_i)$ modelling the relative alignment of the CCD chips then become:

$$\delta x_i = \delta x_S + a_{0i} + a_{1i} \cdot x_{Fi} + a_{2i} \cdot x_{Fi}{}^2$$
$$\delta y_i = \delta y_S + b_{0i} + b_{1i} \cdot x_{Fi} + b_{2i} \cdot x_{Fi}{}^2 \qquad (2)$$

In Eq. 2, i is the index of the CCD chip. The parameters δx_S and δy_S combine corrections for velocity aberration and atmospheric refraction. The constant coefficients a_{0i} and b_{0i} describe the relative shifts of the CCD chips. The coefficient a_{1i} is related to the pixel size, whereas b_{1i} models a shearing of the x_{Fi} axis. The second-order coefficient a_{2i} describes non-linear variations of the pixel size along the x_{Fi} axis, and b_{2i} models a deviation of the shape of the CCD chip from a straight line.

Substituting Eq. 2 into Eq. 1 yields the required modified version of the sensor model for PRISM.

When ALOS PRISM level 1B1 data are imported into *Barista*, one image per CCD chip is instantiated. All the images belonging to the same scene share their exterior orientation and camera mounting parameters. The camera mounting parameters can be found in the metadata files. The exterior orientation parameters, i.e. the coefficients of the spline functions describing the orbit path $\mathbf{S}(t)$ and the rotational angles *roll(t)*, *pitch(t)* and *yaw(t)*, are initialised with approximate values from the orbit observations that are also provided in the metadata [1]. For each CCD chip, one set of alignment coefficients a_{ji} and b_{ji} (cf. Eq. 2) is initialised. For that purpose, the PIs of the ALOS scientific program were provided with information about the relative alignment of the CCD chips inside the camera in the form of camera coordinates for three points per chip (left, centre, and right pixel). If multiple scenes are imported, all sub-scenes recorded by a certain CCD chip will share their alignment coefficients.

There are two ways in which ALOS PRISM level 1B1 imagery can be used for further processing. Firstly, the sub-images from the individual CCD chips can be processed separately, and secondly, the sub-images can be merged to one common image as it would be generated by a single (virtual) camera. In order to produce such a merged image, Eq. 2 and the coefficients of the individual CCD chips are used to model the transformation between the sub-images and the virtual camera. Using the original sub-images has the advantage that the coefficients in Eq. 2 can be determined in adjustment for each CCD chip relative to one master chip, which amounts to a self-calibration of the CCD alignment. Once the sub-images are merged, the link to the original CCD chips is lost, although one set of parameters is maintained that can be used to model remaining systematic errors of the virtual camera. On the other hand, the handling of the sub-images becomes cumbersome, especially in larger projects, and even more so because the two outmost sub-images usually contain only a small strip of meaningful data. Obviously, there is a trade-off between accuracy and productivity. The advantage of using the individual sub-images is only relevant if the initial values of the coefficients in Eq. 2 are considered not to be accurate enough for the application. In this case, they can be determined on-the-job, at the cost of a decreased productivity due to a larger number of GCPs. Otherwise, using the merged images is the better choice.

In [2] we compared bundle block adjustment results using two different sets of parameters for the CCD chip alignment: one early set provided by JAXA in October 2006 and an improved one from July 2007. It was

shown that using the first set of parameters, pixel-level accuracy could be achieved for georeferencing. Using the second parameter set provided by JAXA resulted in sub-pixel accuracy. Though self-calibration could still improve the results in the second case, especially in the height component, our experiments in [2] have shown that JAXA's calibration procedure has been successful. The camera alignment parameters provided in July 2007 are accurate enough to be used for orthoimage production. Hence, we only use merged level 1B1 images in the remainder of this paper.

2.3. The AVNIR-2 Sensor

The interior geometry of the AVNIR-2 level 1 images is different from PRISM [5]. Firstly, each of the four bands of an AVNIR-2 scene (red, green, blue, and near infrared) is provided in a separate image data file, and the sensors recording the individual bands are not perfectly aligned. Secondly, each of the image data files corresponding to one band is generated by the concatenation of images from two different CCD chips that record the even and the odd pixels, respectively. As a consequence, the even and the odd pixel columns of these files are shifted with respect to each other by an offset that depends on the side-looking angle of the sensor. In order to import an AVNIR-2 scene into *Barista*, the side-looking angle has to be read from the metadata files first. It is used to calculate the offset between the even and the odd pixels of the scene. Then each of the image data files corresponding to one band is read and its even pixels are shifted by the computed offset with respect to the odd pixels, which results in a corrected image without the typical stripe pattern that can be observed in the original image data files. In a second stage of preprocessing, the individual bands are merged to a common multi-spectral image. In this context, the bands have to be transformed into the geometry of the green band by a 2D similarity transformation. The parameters of this transformation, especially the constant offsets of the individual bands with respect to the green band, are provided in the metadata files. Theoretically, these parameters could be determined by calibration if GCPs are measured independently in the images corresponding to the individual bands. However, using the parameters found in the metadata files turned out to be accurate enough to deliver multi-spectral images without any visible chromatic distortion.

Once the combined multi-spectral image has been generated for an AVNIR-2 scene, approximate values for its orientation parameters can be determined. We use the same sensor model as for ALOS PRISM, thus a combination of Eqs. 1 and 2. As with merged PRISM level 1B1 scenes, there is only one set of coefficients a_{ji} and b_{ij} per scene that can model systematic errors in the interior orientation parameters. The initial values for the

exterior orientation parameters can be derived from information found in the metadata files in the same way as for PRISM. The camera mounting and the interior orientation parameters can also be derived from information contained in the metadata files.

2.4. Bundle Adjustment

As we have seen in the previous sections, the parameters of the sensor model formulated in Eqs. 1 and 2 can be initialised from information contained in the ALOS metadata files and, in the case of PRISM, from information provided by JAXA. It has been shown that using these data, direct georeferencing is feasible with an accuracy of about 30 pixels [1]. The aim of bundle adjustment is to improve the parameters of the sensor model using the framelet coordinates of GCPs and tie points, the object coordinates of GCPs, and direct observations for the orbit path and attitudes derived from the metadata files. The unknowns to be determined are the coefficients of the spline functions modelling the time-dependant components of the orbit path $\mathbf{S}(t)$ and the time-dependant rotational angles parameterising $\mathbf{R}_P(t)$. Optionally, the coefficients in Eq. 2 can also be determined if self-calibration of the interior camera parameters is to be carried out. The adjustment model is expanded by bias-correction parameters which model the systematic errors in direct observations for the orbit path and attitudes. For each orbit parameter p (the coordinate of an orbit point or a rotational angle), a time-constant unknown Δp is introduced. The observation p^{obs} recorded at time t^{obs} is related to the spline $S_p(t)$ describing the parameter p by:

$$S_p(t^{obs}) = p^{obs} + \Delta p \qquad (3)$$

This results in six parameters for systematic error correction per satellite orbit that have to be determined along with the spline parameters and the coefficients in Eq. 2, these being three offsets $(\Delta X, \Delta Y, \Delta Z)^T$ for the orbit path points and three offsets $(\Delta roll, \Delta pitch, \Delta yaw)^T$ for the rotational angles.

2.5. Strip Adjustment

In the bundle adjustment solution described in the previous section, there is one set of spline parameters and one set of correction parameters per scene. Optionally, there is one set of calibration parameters per camera. Strip adjustment exploits the fact that the satellite orbit is smooth and that the systematic errors in the direct observations for orbit path and attitudes have a long wavelength. If all scenes of a strip refer to the same orbit, there remain only one set of spline parameters and one set of correction parameters to be determined. As a consequence, the number of unknowns is considerably reduced, so that a smaller number of

observations is required to determine these unknowns. If the systematic errors contained in the orbit observations are stable over the length of the strip and if the relative accuracy of these observations is good enough to describe the shape of the orbit, it should be sufficient to use GCPs at the beginning and at the end of the strip. Realisation of the long-strip orientation concept requires an initial merging of the orbit path and attitude data associated with all the scenes of a strip. A single set of exterior orientation, camera mounting, and camera interior orientation parameters applies for the image orientation, and the six bias correction parameters for orbit path and attitude relate to the entire strip.

3. AUTOMATIC DETERMINATION OF GCPs

Even if the GCP requirements can be reduced considerably by strip adjustment, the effort for providing the necessary GCPs in remote areas such as rural Australia can still be prohibitive. If a digital orthoimage collected in the past and a Digital Elevation Model (DEM) are available, the orientation procedure can be fully automated. In this case, one could be excused for thinking that strip adjustment is not required, because the automated procedure could be applied to generate sufficient GCPs for each individual scene. However, in areas such as the Australian deserts it is very difficult to extract features that are visible both in the original orthoimage and the new imagery, so that it is still very useful to reduce the amount of ground control that is required.

The automatic determination of GCPs starts with the definition of areas of interest where GCPs are to be searched. An area of interest is an approximately rectangular area aligned with the strip whose width is equal to the strip width and whose length is equal to a certain percentage (e.g. 33%) of the length of the scene. There are two such areas at the beginning and at the end of the strip and a user-defined number at constant intervals between these two areas. In each of the areas of interest, automatic GCP determination is carried out independently. First, the existing orthoimage covering the area of interest is read, and salient feature points are extracted using the Förstner interest operator [6]. In this context, the user has to select a minimum distance between neighbouring points in the area of interest, in order to control the number of points that is extracted. Typically, this procedure will result in several hundred feature points per area of interest that are candidates for GCPs. These feature points are transformed to object space using the geocoding information of the orthoimage, and the point height is interpolated in the DEM. Thus, the 3D coordinates of the GCP candidates are determined. Their a priori standard deviations in planimetry can be derived from the feature extraction process, whereas the standard deviation of the height component corresponds to the interpolation error in the DEM. Finally, the GCP candidates have to be transformed into the earth-centred coordinate system.

Along with the feature points, small image patches are extracted that are centred at these points. Under the assumption that the area around a feature point is nearly horizontal, the corresponding image patch can be transformed into image space by resampling, using the approximate values of the orientation parameters, the interpolated GCP height, and the geocoding information of the orthoimage. The position of the GCP in image space is then searched by area-based matching, using the transformed image patch as the template and a rectangular area of the satellite image centred at the approximate position of the back-projection of the GCP candidate as the search image. First, cross correlation of the template and the search image is used to determine the position of the GCP candidate with a resolution of 1 pixel [7]. The position of the GCP candidate in the image corresponds to the maximum of the cross correlation coefficient between template and search image. If this maximum is below a user defined threshold, e.g. 0.8, the candidate is discarded. In this case, the orthoimage around the extracted feature point is no longer similar to the satellite image. Otherwise, fine measurement of the position is carried out by Least Squares Matching [8].

All GCP candidates that are successfully matched in one of the satellite images are used for strip adjustment. This results in a highly redundant adjustment, and the high redundancy can be used to eliminate remaining false matches by robust estimation [9].
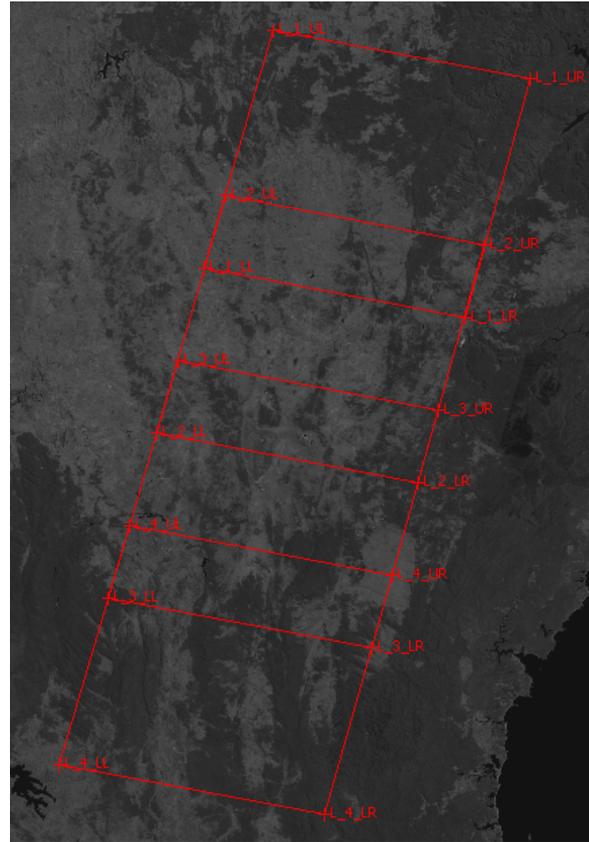
4. EXPERIMENTS

4.1. Test Data

In order to validate the newly developed long-strip image orientation process for ALOS PRISM and AVNIR-2 imagery, an initial experimental evaluation was performed on two sets of images. The first was a strip of 10 overlapping PRISM scenes covering a length of 293 km with a width of 37 km, centred over Canberra, Australia, as shown in Fig. 1a. The second, Fig. 1b, was a strip of four overlapping AVNIR-2 scenes over the same area, which covered 250 km in length by 72 km in width. Some 123 3D points were available for use as either control or checkpoints. These points were determined by GPS with an accuracy of about 1 m in both planimetry and height. Nearly all points were determined at road intersections. Since most of the roads in the area are unsealed, the definition of these points is not always optimal. In Fig. 1a, the available 3D points are displayed in yellow.
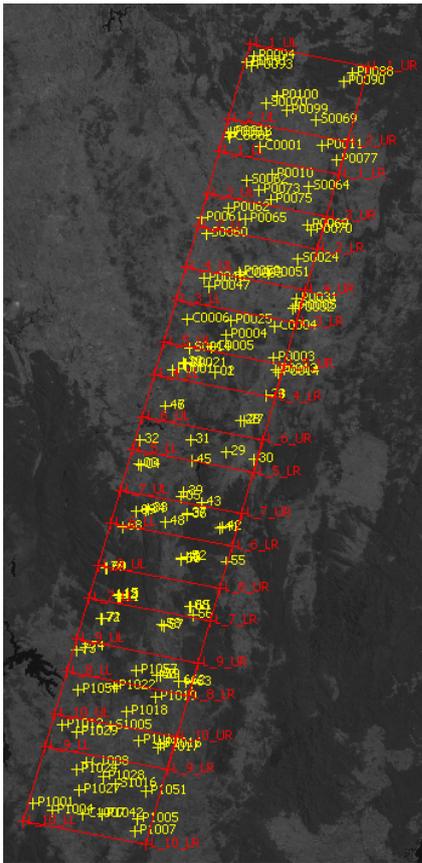
(a) PRISM   *(b) ANIR-2*

Figure 1. Strips of ALOS imagery for experimental testing of strip orientation.

In order to test how long a strip can be, the PRISM strip shown in Fig. 1a was expanded both to the north and to the south by additional 11 scenes, which resulted in a strip of altogether 21 scenes with a length of about 600 km. Eleven and five additional 3D points were available in the southernmost and northernmost scenes, respectively. The 10 scenes of the first test were situated in the centre of the extended strip, which is where the maximum errors would be expected in strip adjustment. For economical reasons, no additional check points could be determined in the scenes between beginning / end of the long strip and the 10-scene strip.

For automatic GCP determination, an orthoimage mosaic of Australia generated from Landsat panchromatic scenes at a ground resolution of 0.45" is available at GA. Where no other information can be obtained, especially in the Australian outback, this image will be used to define GCPs at GA. Thus, the resolution of the orthoimage is below the resolution of PRISM approximately by a factor of 5; its resolution is still slightly worse than the resolution of AVNIR-2. One of the only two nationwide DEMs available is the SRTM DEM, with a height accuracy of about 10 m. For the tests in this paper, we could use a state-wide DEM

generated by the Department of Lands of New South Wales, which has a better accuracy. Of course, using ground control that is worse in its accuracy than the PRISM resolution is not an optimal solution; it can be justified by economical reasons and by the fact that it will only be used if no other information is available. Wherever better imagery is available for GCP definition, e.g. if it is provided by a customer of GA, it will be used for that purpose.

4.2. Validation on the 10-scene PRISM strip

Shown in Tab. 1 are the RMS errors of the coordinate differences at check points in image space that resulted from the image orientation of the 10-scene PRISM strip for four different variants of adjustment. In the first scenario, all satellite orbits were determined independently from each other, using four GCPs in each scene. Since there is a considerable overlap between the scenes, some GCPs are used in two images. The total number of GCPs used is 23. The results are consistent with those reported in [2], with overall RMS errors of about 0.65 pixels both in X and Y. It is obvious that sub-pixel accuracy can be achieved.

Table 1. RMS errors at check points in image space [pixels] from bundle adjustment of 10 PRISM scenes. N is the number of check points used in the scene.

Scene	Separate orbits 23 GCPs			Single orbit 8 GCPs			Single orbit 4 GCPs			Single orbit 0 GCPs (matching)		
	N	RMS_X	RMS_Y	N	RMS_X	RMS_Y	N	RMS_X	RMS_Y	N	RMS_X	RMS_Y
P_1	11	0.67	0.60	11	0.64	0.58	13	0.63	0.54	15	1.15	2.00
P_2	12	0.52	0.41	14	0.52	0.38	16	0.53	0.40	16	0.95	1.78
P_3	10	0.89	0.77	14	0.62	0.61	14	0.51	0.76	14	1.16	1.96
P_4	15	0.68	0.69	19	0.55	0.68	19	0.49	1.01	19	1.09	2.28
P_5	15	0.59	0.56	19	0.49	0.54	19	0.49	0.67	19	1.03	1.85
P_6	14	0.89	0.51	18	0.57	0.52	18	0.66	0.51	18	1.32	1.72
P_7	19	0.72	0.66	23	0.57	0.68	23	0.62	0.73	23	1.27	1.86
P_8	17	0.67	0.75	21	0.56	0.76	21	0.57	0.88	21	1.14	2.01
P_9	14	0.49	0.55	16	0.59	0.64	18	0.57	0.71	18	1.14	1.82
P_{10}	14	0.49	0.74	14	0.48	0.60	16	0.58	0.91	18	1.29	1.93
Total	**141**	**0.67**	**0.64**	**169**	**0.56**	**0.62**	**177**	**0.57**	**0.74**	**181**	**1.16**	**1.93**

In all the other scenarios, the orbits of all scenes were merged, so that only one set of spline parameters and one set of error correction parameters had to be determined for the orbit path and attitudes, respectively. No correction parameters for systematic errors in the camera (cf. Eq. 2) were determined. First, we reduced the set of GCPs to those that were used in the southernmost and northernmost scenes. The results achieved for that scenario are nearly identical or even slightly better than those achieved in the adjustment with separate orbits, and sub-pixel accuracy can still be achieved. When the number of GCPs was reduced further and only four were used in the corners of the strip, the adjustment results were still quite good, with an accuracy at pixel-level or better. The most important result thus obtained was that there is little distinction in accuracy between the three cases. The results of the separate and single orbit cases produce the same overall RMS coordinate error in planimetric accuracy of 0.66 pixels, though there is marginally better homogeneity of accuracy between the along- and cross-track directions for the separate orbit case. The absolute RMS accuracy, as quantified by checkpoints, is very impressive, being about 1.7 m, with the maximum coordinate discrepancy being 5.5 m or 2.2 pixels. The reason that there are more than 123 checkpoints indicated in each list is that the checkpoint discrepancies were computed for each scene and the scenes overlap.

In the last scenario of adjustment, no GPS-surveyed GCPs were used. Instead, ground control was determined automatically through image matching in the way described in Section 3. Three GCP areas were defined in the strip. Altogether 8961 features were extracted in these areas. The size of the image patches for matching was 31 pixels. A point was only accepted if the cross correlation coefficient was greater than 0.9. This resulted in 205 3D points that could be used as GCPs. Fig. 2 shows the distribution of these points over the PRISM strip. The small number of GCPs in southernmost scene of the strip is due to the fact that the
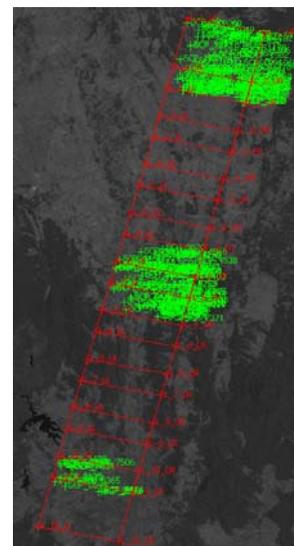

Figure 2. Distribution of the GCPs determined automatically in the 10 PRISM scenes.

DEM only covered the northern part of that scene. In the course of strip adjustment, one point was eliminated as a gross error.

The results of the error assessment in Tab. 1 show that in this case, not even pixel-level accuracy could be achieved. The RMS values vary between 1 and 2.3 pixels. The maximum difference is 3.6 pixels. The automated process results in a deterioration of the accuracy by a factor 2-3 compared to a scenario where GCPs determined by GPS are used. The errors are much larger in Y than in X. There is a considerable systematic shift of 1.0 pixel in X and 1.8 pixels in Y. These values correspond to the resolution of PRISM; in the resolution of the orthoimage, all RMS errors are smaller than 0.4 pixels. The accuracy of the automated process is clearly limited by the accuracy of georeferencing and the resolution of the orthoimage. The bias in the Y coordinates indicates a systematic shift of about 1/3 pixel in the georeferencing of the orthoimage.

Table 2. RMS errors at check points in image space [pixels] from bundle adjustment of 4 AVNIR scenes. N is the number of check points used in the scene.

Scene	Separate orbits 11 GCPs			Single orbit 4 GCPs			Single orbit 0 GCPs (matching)		
	N	RMS_X	RMS_Y	N	RMS_X	RMS_Y	N	RMS_X	RMS_Y
A_1	14	0.36	0.56	16	0.38	0.61	18	0.42	0.52
A_2	23	0.34	0.47	27	0.34	0.58	27	0.44	0.99
A_3	23	0.27	0.45	27	0.25	0.57	27	0.46	0.78
A_4	28	0.39	0.45	31	0.38	0.62	33	0.56	0.97
Total	**88**	**0.34**	**0.47**	**101**	**0.34**	**0.59**	**105**	**0.48**	**0.87**

4.3. Validation on AVNIR-2

For AVNIR-2, three different adjustments were carried out. The results are listed in Tab. 2. For AVNIR-2 it turned out to be advantageous to determine systematic error correction parameters for the camera (cf. Eq. 2) in all scenarios. Again, we started with an adjustment using separate orbits and four GCPs per scene. In this case, 11 GCPs were used. The RMS values are all below 0.6 pixels, with the RMS error in Y (i.e., in flying direction) being systematically larger than the RMS error in X. In the second scenario, we applied strip adjustment, and the number of GCPs was reduced to four in the corners of the strip. The results in Tab. 2 indicate that again a slightly larger RMS error was produced in the along-track direction in the single-orbit adjustment, though the increase in the RMS Y value from 0.5 pixels to 0.6 pixels was of no practical significance. Once again, sub-pixel RMS accuracies were obtained in the strip adjustment, these being 0.34 pixels (3.4 m) in the cross-track direction and 0.6 pixels (6.0 m) in the along-track direction. The largest of the 101 checkpoint residuals was 1.5 pixels.

For automatic GCP determination, two GCP areas were defined. Altogether 552 features were extracted in these areas. The size of the image patches for matching was selected to be 35 orthoimage pixels. A matched point was only accepted if the cross correlation coefficient was greater than 0.8. This resulted in 148 3D points that could be used as GCPs. Fig. 3 shows the distribution of these points over the AVNIR-2 strip. There were no gross errors.

The results of error assessment in Tab. 2 show that the RMS values vary between 0.4 and 1.0 pixels. The maximum difference is 1.8 pixels. The automated process results in a deterioration of the accuracy by a factor 2 compared to a scenario where well-defined GCPs determined by GPS are used. The errors are larger in Y than in X. There is also a systematic shift of 0.4 pixels in X and 0.75 pixels in Y. This systematic shift in Y corresponds to about 0.5 pixels in the digital orthoimage. These results show that pixel-level accuracy can be achieved for the georeferencing of AVNIR-2 strips using the automated measuring tool,
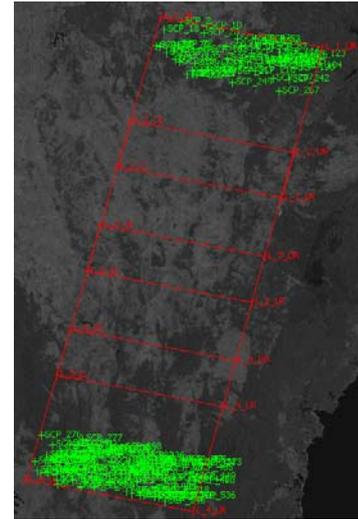

Figure 3. Distribution of the GCPs determined automatically in the 4 AVNIR-2 scenes.

but the results cannot be better than the georeferencing of the orthoimage used to extract the GCPs.

4.4. Validation on the Long PRISM Strip

The results described in Section 4.2 show that by combining 10 PRISM scenes the potential of strip adjustment is not fully exploited. Using the expanded strip, two more scenarios of strip adjustment were tested, in both cases with a single orbit. In the first scenario, we used 8 GCPs (four in the northernmost and southernmost scenes), and in the second scenario we used only four GCPs, i.e. one per corner of the strip.

The results of error assessment are presented in Tab. 3. Note that check points were only available in the centre of the strip. This is where the largest deviations would be expected. Our results show that in both scenarios, the RMS errors are all below one pixel. Using 8 GCPs resulted in RMS errors that are almost identical to those achieved by an adjustment using separate orbits and four GCPs per scene (cf. Tab. 1). If the number of GCPs is reduced further, the results achieved are still quite good, though in this case the RMS values in X are slightly worse and can reach 0.9 pixels. The overall

RMS values of 0.6 and 0.7 pixels (corresponding to 1.5 m and 1.8 m, respectively) are still very impressive. Assuming that in the case of an adjustment using separate orbits with four GCPs per scene the GCPs in the overlap areas would be measured in two images, a total number of 44 GCPs would be required for precise georeferencing of the 21 scenes. Thus, our results indicate that the number of GCPs can be reduced by 80% without any loss in accuracy and even by 90% if a slight deterioration of the accuracy is accepted.

Table 3. RMS errors at check points in image space [pixels] from adjustment of 21 PRISM scenes. N is the number of check points used in the scene.

Scene	N	8 GCPs		4 GCPs	
		RMS_X	RMS_Y	RMS_X	RMS_Y
P_1	15	0.68	0.49	0.85	0.49
P_2	16	0.68	0.33	0.90	0.34
P_3	14	0.55	0.68	0.71	0.68
P_4	19	0.56	0.82	0.75	0.74
P_5	19	0.59	0.57	0.81	0.57
P_6	18	0.55	0.45	0.64	0.50
P_7	23	0.53	0.63	0.65	0.66
P_8	21	0.55	0.80	0.73	0.83
P_9	18	0.58	0.61	0.72	0.73
P_{10}	18	0.56	0.65	0.61	0.71
Total	**181**	**0.58**	**0.63**	**0.74**	**0.68**

5. CONCLUSION

This paper has overviewed the development and testing of an image orientation process for long strips of ALOS PRISM and AVNIR-2 imagery. Through the adoption of a rigorous sensor orientation model, successive scenes of imagery recorded on a single pass of the satellite can be merged. By this means, drastic reductions in ground control are possible in the georeferencing and subsequent orthoimage generation, without any associated loss in accuracy. An automatic technique for GCP definition and measurement was also implemented and tested, and it was shown that the results that can be achieved by the automated process are mainly limited by the quality of the data that are used to automatically define GCP candidate points. For the data used in our test, pixel-level results could be achieved for AVNIR-2 by the fully automated work flow. For PRISM, whose resolution is better than the resolution of the orthoimage used for GCP definition by a factor of 5, the RMS errors achieved by the fully automatic work flow were in the range of two pixels.

Following experimental testing that has indicated that sub-pixel geopositioning accuracy can be achieved for PRISM strips of 600 km in length, comprising 20 or more scenes, with only four to eight GCPs, the strip adjustment process has been implemented for the production of orthoimages from ALOS imagery at Geoscience Australia. Initial indications are that largely

due to the significant reduction in required ground control and the ability to process multi-scene strips instead of single scenes, the rate of production has been enhanced by more than 300%.

6. REFERENCES

1. Weser, T., Rottensteiner, F., Willneff, J., Poon, J. & Fraser, C. (2008). Development and testing of a generic sensor model for high-resolution satellite imagery. *Photogrammetric Record* **21**(123), 255-274.

2. Weser, T., Rottensteiner, F., Willneff, J. & Fraser, C. S. (2008). An improved pushbroom scanner model for precise georeferencing of ALOS PRISM imagery. *Intern. Arch. Photogrammetry, Remote Sensing & Spatial Information Sci.*, Beijing, XXXVII (B1-2), pp 723-730.

3. Barista product information webpage (2008). *http://www.baristasoftware.com.au* (Accessed: 14th October 2008).

4. JAXA (2006). ALOS PRISM Level 1 Product Format Descriptions Rev. J. Available via *http://www.eorc.jaxa.jp/ALOS/doc/format.htm* (Accessed: 14th October 2008).

5. JAXA (2006). ALOS AVNIR-2 Level 1 Product Format Descriptions Rev. J. Available via http://www.eorc.jaxa.jp/ALOS/doc/format.htm (Accessed: 14th October 2008).

6. Förstner W. & Gülch, E. (1987). A fast operator for detection and precise location of distinct point, corners and centres of circular features. Proc. ISPRS Conference on Fast Processing of Photogrammetric Data, Interlaken, pp.281-305.

7. Paar, G., Rottensteiner, F. & Pölzleitner, W. (2001). Image Matching Strategies. Kropatsch, W. & Bischof, H. (eds), *Digital Image Analysis. Selected Techniques and Applications*, Springer, New York, USA, pp 393-401.

8. Baltsavias, E. (1991). *Multiphoto Geometrically Constrained Matching.* PhD Thesis, Institute of Geodesy and Photogrammetry, ETH Zurich, Switzerland, Mitteilungen Nr 49.

9. Kager, H. (1989). ORIENT: A universal photogrammetric adjustment system. Grün, A. & Kahmen, H. (eds), *Optical 3-D Measurement*, H. Wichmann, Karlsruhe, Germany, pp 447-455.