# IMAGE MERGING TO SUPPORT GEOREFERENCING AND ORTHOIMAGE GENERATION FROM ALOS IMAGERY

C.S. Fraser & T. Weser
Cooperative Research Centre for Spatial Information, Department of Geomatics
University of Melbourne, VIC 3010, Australia
Email: {c.fraser, tweser}@unimelb.edu.au

F. Rottensteiner
Institute of Photogrammetry and Geoinformation
Leibniz University of Hanover, D-30167 Hannover, Germany
Email: rottensteiner@ipi.uni-hannover.de

**KEY WORDS:** ALOS, PRISM, AVNIR-2, sensor orientation, strip adjustment, orthoimages

**ABSTRACT:** Imagery from the ALOS PRISM and AVNIR-2 sensors offers the potential of georeferencing and orthoimage generation to support medium- and small-scale mapping. In order to optimize the economy and productivity of these operations, image strips of maximum possible length need to be employed with a minimum of ground control, without suffering a loss in attainable accuracy. This paper describes the development of a computational system that enables realisation of this goal. The data processing flow commences with an image merging stage that results in a single scene for AVNIR and three single scenes for PRISM. In the case of PRISM, Level 1B1 sub-images are merged, and for AVNIR individual image bands are merged. Along with the merging of imagery, all sensor orientation data is rigorously transformed and referenced to a single orbit, with one set of orbit path and attitude parameters. A strip adjustment is then carried out to refine the sensor orientation. This employs four or more ground control points at the two end regions of the strip. The refined sensor orientation modelling accounts for biases within the orbit and attitude data, resulting in 1-pixel and even sub-pixel level georeferencing accuracy for all scenes of the strip. This paper overviews the computational models and steps involved, for both PRISM and AVNIR-2 imagery, and describes how the workflow has been accomplished within the *Barista* software system. Practical experimental test results are presented, in which it is demonstrated that 1-pixel georeferencing accuracy and better is achievable for PRISM imagery with as few as four ground control points over strip lengths of 700km. Such a control requirement is well below that generally needed for orthoimage generation approaches that neither employ rigorous sensor orientation modelling for ALOS imagery nor have the capability to merge along-track scenes into long continuous strips.

## 1. INTRODUCTION

Two of the image products offered by Geoscience Australia (GA) are orthoimages generated from ALOS PRISM and AVNIR-2 imagery. The production sequence, indicated in Fig. 1, has involved the preprocessing of imagery to 1B2 level, the provision of ground control points (GCPs), the orientation for every image scene, and the generation of orthoimages using an SRTM DEM. One of the constraints to production has been the requirement to produce the orthoimages scene-by-scene, which is time intensive, especially in regard to providing the necessary 10+ GCPs for each image, even though this process can be automated via image matching utilizing a database of control point chips. Productivity in orthoimage generation 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 1000 $km^2$.

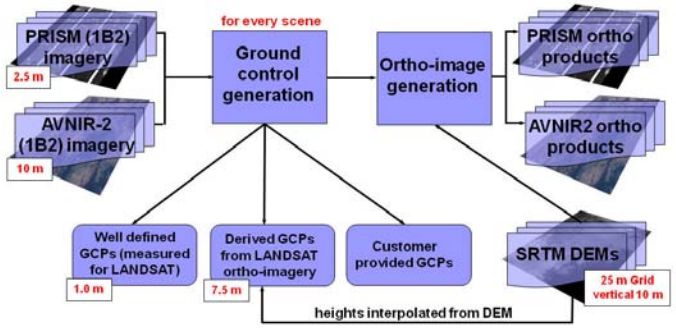



Figure 1: Production of ALOS orthoimages at GA using single scenes.

One of the ways to overcome the GCP requirements is to perform multi-scene processing, but such a capability was not offered by the commercial photogrammetric data processing system used by GA. Thus, the Cooperative Research Centre for Spatial Information (CRCSI) was approached to ascertain whether this functionality could be developed within the *Barista* software system for satellite image processing (Barista, 2008). This paper describes such a development, which required the merging of successive ALOS PRISM and AVNIR-2 scenes from a single orbit into single long strips of imagery. As will be seen, this then allows the rigorous orientation of strips formed by 20+ scenes, using as few as four GCPs positioned at the corners of the strip. The scenario for the production of orthoimages then becomes that shown in Fig. 2, where the chief practical distinction from Fig. 1 is that far fewer GCPs are required for the sensor orientation, with the image now covering 500-700km rather than 35km, for example.

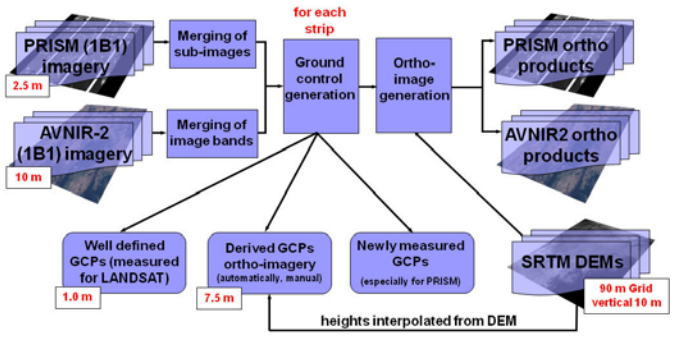



Figure 2: Production of orthoimages using long strips of merged ALOS scenes.

In the following sections, a brief description of the generic sensor orientation model (Weser et al, 2008a;b) is presented, along with specific requirements for the orientation of both ALOS PRISM and AVNIR-2 imagery via this model. This is followed by a brief description of the process by which successive scenes are merged into a single 'long strip' scene to support bridging without ground control. An account of the results of experimental application of the long-strip image orientation is then presented.

## 2. SENSOR ORIENTATION MODEL FOR ALOS STRIP-IMAGES

### 2.1 Generic Orientation Model

The physical model of the pushbrooom 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 directly 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 $\Delta t$ the time interval for recording a single image row. The framelet coordinate system refers to an individual CCD array. In that coordinate system an image observation 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 of the earth's surface recorded at time $t$ that corresponds to $y_I$. The relation between an observed image point $\mathbf{p}_F$ in and the object point $\mathbf{P}_{ECS}$ is described by Eq. 1:

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

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, each image 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. The rotation matrix $\mathbf{R}_O$ acts as an angular offset. It 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 on the imaging mode, four or six CCD chips are used to record a scene. 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 (see Weser et al., 2008b). 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:

$$\begin{aligned} \delta x_i &= \delta x_S + a_{0i} + a_{1i} \cdot x_{Fi} + a_{2i} \cdot x_{Fi}{}^2 \\ \delta y_i &= \delta y_S + b_{0i} + b_{1i} \cdot x_{Fi} + b_{2i} \cdot x_{Fi}{}^2 \end{aligned} \quad (2)$$

In Eq. 2, $i$ is the index of the CCD chip. The parameters $\delta x_S$ and $\delta y_S$ 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 ALOS PRISM.

### 2.3 The AVNIR-2 Sensor

Whereas PRISM sub-images are first merged into a common camera and framelet system for image orientation, the requirement for AVNIR-2 centres upon the need to preprocess the separate spectral bands of Level 1A and 1B1 imagery to remove band misalignment. This can be achieved through the use of additional parameters that model the required shift, differential scale change and rotation between the bands. These parameters are provided with sufficient accuracy in the metadata files.

### 2.4 Bundle Adjustment

The aim of bundle adjustment is to improve the parameters of the sensor model formulated in Eqs. 1 and 2 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 coefficients of the spline functions modelling the time-dependant components of the orbit path $\mathbf{S}(t)$ and the time-dependant rotational angles parameterising $\mathbf{R}_P(t)$ are also determined. 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 $\Delta p$ is introduced. The observation $p^{obs}$ recorded at time $t^{obs}$ is related to the spline $S_p(t)$ describing the parameter $p$ by:

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

This results in six parameters for systematic error correction per satellite orbit that have to be determined along with the spline parameters, 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

Realisation of the long-strip orientation concept, illustrated by Fig. 3, requires an initial merging of scenes along with their associated orbit path and attitude data. A single set of camera mounting and camera interior orientation parameters then applies for the image orientation, and the six bias correction parameters for orbit path and attitude relate to the entire multi-scene image. The resulting adjusted orientation and bias-correction parameters for the single long-strip can then be mapped back to the individual scenes to support single scene orthoimage generation.

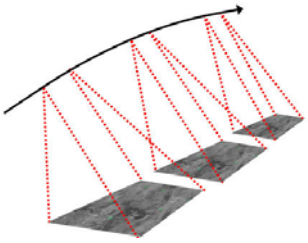


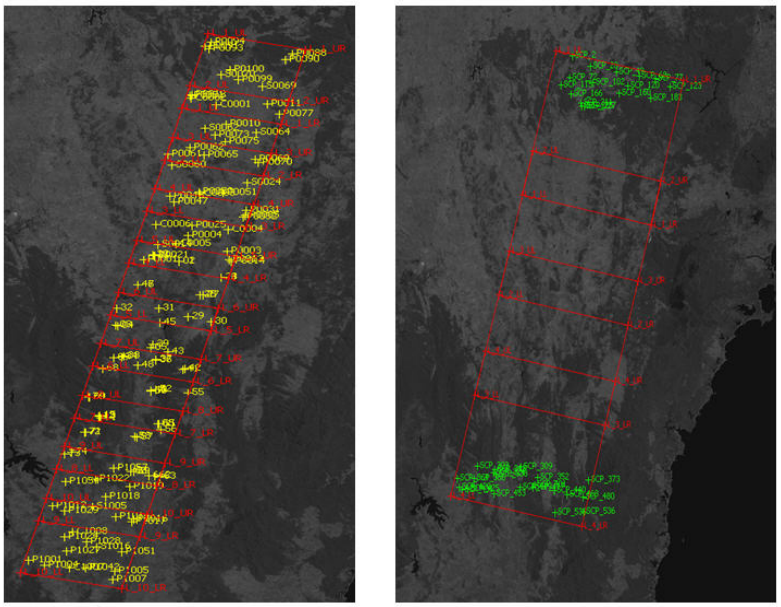
Figure 3: Merging of single ALOS scenes into a long-strip for image orientation.

## 3. EXPERIMENTAL RESULTS

### 3.1 Initial Model Validation Testing on PRISM and AVNIR-2

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 37km, centred over Canberra, Australia, as shown in Fig. 4a. The second, Fig. 4b, 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 GCPs were available for use as either control or checkpoints. These points had a nominal sub-metre accuracy in both planimetry and height.



(a) PRISM (b) AVNIR-2

Figure 4. Strips of ALOS imagery for experimental testing of strip orientation.

Shown in Table 1 are the RMS errors in planimetry that resulted from the image orientation of the 10-scene PRISM strip when adjusted both scene by scene (separate orbits) with 4 corner GCPs (Table 1a), and as a single strip image (single orbit), again with 4 corner GCPs (Table 1b). The most important result obtained was that there is little distinction in accuracy between the two 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. It is noteworthy that 23 GCPs were needed for adjustment of the scenes as separate orbits, whereas only 4 GCPs were required for the single-orbit case. The absolute RMS accuracy, as quantified by checkpoints, is very impressive, being 1.7m, with the maximum coordinate discrepancy being 5.5m 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.

| Scene | Points | RMSX | RMSY | MaxX | MaxY |
|---|---|---|---|---|---|
| 1 | 11 | 0.67 | 0.60 | 1.51 | 1.31 |
| 2 | 12 | 0.52 | 0.41 | 0.96 | 0.83 |
| 3 | 10 | 0.89 | 0.77 | 1.50 | 1.80 |
| 4 | 15 | 0.68 | 0.69 | 1.59 | 1.63 |
| 5 | 15 | 0.59 | 0.56 | 1.52 | 1.12 |
| 6 | 14 | 0.89 | 0.51 | 1.40 | 0.99 |
| 7 | 19 | 0.72 | 0.66 | 1.43 | 1.44 |
| 8 | 17 | 0.67 | 0.75 | 1.29 | 1.51 |
| 9 | 14 | 0.49 | 0.55 | 0.93 | 1.05 |
| 10 | 14 | 0.49 | 0.74 | 1.37 | 1.63 |
| **Total** | **141** | **0.67** | **0.64** | **1.59** | **1.80** |

(a)

| Scene | Points | RMSX | RMSY | MaxX | MaxY |
|---|---|---|---|---|---|
| 1 | 13 | 0.63 | 0.54 | 1.29 | 1.04 |
| 2 | 16 | 0.53 | 0.40 | 1.01 | 0.87 |
| 3 | 14 | 0.51 | 0.76 | 0.89 | 2.05 |
| 4 | 19 | 0.49 | 1.01 | 0.98 | 2.18 |
| 5 | 19 | 0.49 | 0.67 | 1.26 | 1.51 |
| 6 | 18 | 0.66 | 0.51 | 1.17 | 1.22 |
| 7 | 23 | 0.62 | 0.73 | 1.41 | 1.91 |
| 8 | 21 | 0.57 | 0.88 | 1.18 | 1.71 |
| 9 | 18 | 0.57 | 0.71 | 1.15 | 1.50 |
| 10 | 16 | 0.58 | 0.91 | 1.17 | 1.75 |
| **Total** | **177** | **0.57** | **0.74** | **1.41** | **2.18** |

(b)

Table 1. RMS errors in planimetry (in pixels) from bundle adjustment of multi-scene ALOS PRISM imagery, as (a) separate orbits and (b) a single orbit for the 10 scenes.

The corresponding results for AVNIR-2, listed in Tables 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 from 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 (3.4m) in the cross-track direction and 0.6 pixels (6m) in the along-track direction. The largest of the 100 checkpoint residuals was 1.5 pixels.

| Scene | Points | RMSX | RMSY | MaxX | MaxY |
|---|---|---|---|---|---|
| 1 | 14 | 0.36 | 0.56 | 0.97 | 1.00 |
| 2 | 23 | 0.34 | 0.47 | 0.86 | 0.88 |
| 3 | 23 | 0.27 | 0.45 | 0.71 | 1.00 |
| 4 | 28 | 0.39 | 0.45 | 1.49 | 1.13 |
| **Total** | **88** | **0.34** | **0.47** | **1.49** | **1.13** |

| Scene | Points | RMSX | RMSY | MaxX | MaxY |
|---|---|---|---|---|---|
| 1 | 16 | 0.38 | 0.61 | 1.10 | 1.22 |
| 2 | 27 | 0.34 | 0.58 | 1.03 | 1.10 |
| 3 | 27 | 0.25 | 0.57 | 0.65 | 1.04 |
| 4 | 31 | 0.38 | 0.62 | 1.47 | 1.40 |
| **Total** | **101** | **0.34** | **0.59** | **1.47** | **1.40** |

Table 2. RMS errors in planimetry (in pixels) from bundle adjustment of multi-scene ALOS AVNIR-2 imagery, as (a) separate orbits and (b) a single orbit for the 4 scenes.

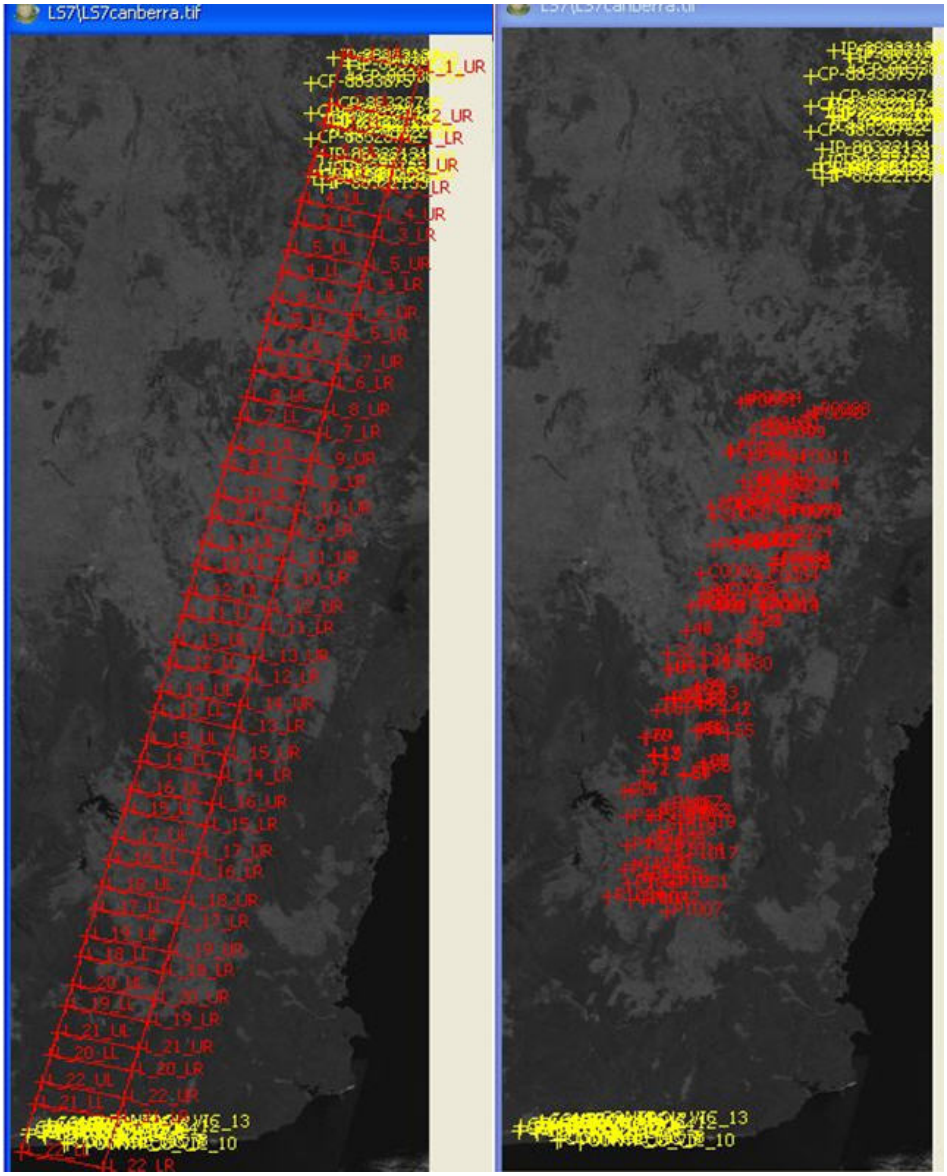
### 3.2 22-Scene PRISM Strip

As a follow-up to the initial experimental testing of the 10-scene PRISM and 4-scene AVNIR-2 strips, further tests were conducted on the 22-scene PRISM strip shown in Fig. 5. At this writing (October 2008), the full series of tests for this strip of PRISM images, which is an extension in length of the image sequence described above, has not been completed, but what can be stated is that the adjustment of this 700km long strip using a single orbit produces sub-pixel checkpoint RMS accuracies when as few as 6 GCPs are employed, nominally 3 at each end. Moreover, as indicated in Fig. 5b, most of the checkpoints are clustered in the centre of the strip where results would be expected to be the worst. Yet, for points determined in the bundle adjustment with bias correction, RMS errors in planimetry of 1.6m were obtained from 50 checkpoints. Thus, optimal image orientation accuracy is maintained in the strip adjustment, while the GCP requirement is drastically reduced, by a factor of around 10 in this case.

## 4. CONCLUDING REMARKS

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, effectively into a single long-strip image where the parameters of the image orientation relate to a single orbit. By this means, drastic reductions in ground control are possible in the georeferencing and subsequent orthoimage generation, without any associated loss in accuracy. Following experimental testing that has indicated that sub-pixel geopositioning accuracy can be achieved for PRISM strips of 700km in length, comprising 20+ scenes, with only six GCPs and possibly only four, 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%. Work is continuing on the assessment of the performance of the new strip adjustment model, which is implemented in the *Barista* software system. The process will next be tested on other high resolution satellite imaging

systems, such as Quickbird and WorldView-1, that provide orbit path and attitude information as part of the metadata.




(a) Scene layout (b) Available GCPs & checkpoints

Figure 5. 22-Scene, 700km Strip of ALOS PRISM imagery.

## 5. REFERENCES

Barista, 2008. Barista product information webpage, *http://www.baristasoftware.com.au* (Accessed: 21. April 2008).

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

Weser, T., Rottensteiner, F., Willneff, J. and Fraser, C.S. , 2008b. An Improved Pushbroom Scanner Model for Precise Georeferencing of ALOS PRISM Imagery. *Intern. Archives of Photogrammetry, Remote Sensing and Spatial information Sciences*, Beijing, XXXVII(B1-2): 723-730.