RIGOROUS IMAGE FORMATION FROM AIRBORNE AND SPACEBORNE DIGITAL ARRAY SCANNERS

H. J. Theiss

InnoVision Basic and Applied Research Office, Sensor Geopositioning Center, National Geospatial-Intelligence Agency (contractor) PA Case 12-189, 7500 GEOINT Dr, Springfield, VA 22150, USA – Henry.J.Theiss.ctr@nga.mil

Commission I, WG 3

KEY WORDS: geometric, modelling, optical, orientation, processing, rectification, restitution, scanner

ABSTRACT:

Sensor builders in the digital era have design limitations due to the constraint of maximum available digital array size. A straightforward solution exists, for example, when four cameras that each simultaneously captures an image from essentially the same perspective centre; they can be re-sampled to form a virtual large format image that can be exploited using a single (instead of four separate) instantiation of a frame model. The purpose of this paper is to address the less trivial time-dependent cases where the sensor scans the ground and the detector arrays obtain chips of imagery that need to be stitched together to form a single conveniently exploitable image. Many operational techniques warp the imagery to form a mosaic, or ortho-rectify it using an imperfect digital surface model (DSM), thus eliminating the possibility for accurate geolocation and uncertainty estimation. This algorithm, however, forms a single virtual image with associated smooth metadata, which can be exploited using a simple physical sensor model. The algorithm consists of four main steps: 1) automated tie point matching; 2) camera calibration (once per sensor); 3) block adjustment; and 4) pixel re-sampling based on an "idealized" virtual model. The same geometry model used to form the image, or its true replacement, must be used to exploit it. This paper verifies the algorithm using real imagery acquired from the Global Hawk (GH) UAV. Registration of the virtual image to a WorldView1 stereopair using four tie points yielded an RMS below 0.6 meters per horizontal axis.

1. INTRODUCTION

1.1 Problem Statement

Image acquisition in the era of film recording fell primarily into three main classes: frame, strip, and panoramic. For each of these classes the recorded image was identical to the image exploited by the user since no resampling of pixels was required to form the image. Hence, the sensor model used to perform photogrammetric operations corresponded to physical parameters associated with the original recorded image. As we entered into the digital age, sensor builders encountered challenges due to the maximum size digital array that could be manufactured.   An example of how this limitation was overcome is the DMC® from Z/I Imaging® which simultaneously acquires four high resolution panchromatic images from essentially the same perspective centre, and re-samples the pixels to form a single virtual large format image. Photogrammetric exploitation of the resulting image is relatively simple in that a single frame sensor model is used instead of four separate instantiations of a frame model. The purpose of this paper is to address the less trivial time-dependent cases, such as pushbroom or whiskbroom, where the sensor scans the ground and the detector arrays obtain chips of imagery that need to be stitched together to form the image.

1.2 Overview of Approach

This paper provides an algorithm and photogrammetric modelling results for forming a single image, which can be exploited using a simple generic physical sensor model, given several image chips collected over time. One data set used to verify the algorithm is real imagery acquired from the Global Hawk (GH) UAV. It scans a square digital frame array obtaining 14 cross-track scans of 10 frames each.   The algorithm forms a single virtual image with associated smooth metadata, which can be exploited using a simple generic whiskbroom sensor model.   Thorough metric analyses have been performed on these virtual image products to demonstrate no loss of precision in the derived coordinates relative to the reference imagery, i.e. a stereo pair of WorldView1 images with 0.5 meter GSD. The motivation for implementing this new image formation approach is that many operational techniques warp the imagery to form a mosaic thus eliminating the possibility for accurate geolocation and uncertainty estimation. The only remaining option for rigorous exploitation had been to exploit the 140 frames independently, but users had found it too cumbersome.

The algorithm, which can be applied to most any time dependent imaging system besides GH, consists of four main steps: 1) obtain tie point correspondences in the narrow forward- and side-lapping regions of the frames; 2) perform camera calibration; 3) perform block adjustment; and 4) re-sample the pixels to obtain a single virtual image with associated smoothed metadata.   The first step, tie point matching, is performed using normalized cross correlation. The camera calibration is performed only once for the sensor, and the recovered values are used to form subsequent images. The block adjustment solves for corrections to roll, pitch, and yaw angles for each frame, and a single set of three translations for the entire block.   Tie points are free to move in the two directions perpendicular to the line of sight, while they are constrained in the line-of-sight direction by an amount that is a function of the convergence angle between the rays. A post-adjustment range map across the scene is generated as a function of the tie point 3D locations for use in the next step. Prior to performing the actual re-sampling, the position and

attitude data are smoothed by fitting low order polynomials. To form the final virtual image, the algorithm requires that the following steps be performed per pixel in the virtual image. First, use the generic whiskbroom sensor model associated with the virtual image, its smoothed metadata, and the range map to perform an image-to-ground (i2g) transformation. Given this ground point, perform a ground-to-image transformation into the nearest frame using the frame sensor model and its associated interior orientation parameters obtained by the camera calibration. Finally, bi-linearly interpolate to obtain the gray value to assign to the virtual image pixel.

In order to exploit the virtual image, the same geometry model must be used that was employed to perform the i2g step of re-sampling associated with the image formation. After registering the virtual image to a WorldView1 stereopair using 4 tie points, comparison with over 140 check points demonstrated an RMS below 0.6 meters per horizontal axis which was essentially equivalent to the accuracy of the WorldView1 derived check points with respect to a local datum. Notably, the rigorous image formation approach provides over an order of magnitude accuracy improvement compared to the operational image warping approach.

1.3 Paper Outline

Following this introduction section, the paper proceeds with Section 2, a description of the real airborne data set used to demonstrate the success of this image formation algorithm. Then Section 3 provides a high level overview of the automated tie point matching. Sections 4 and 5 discuss the camera calibration and bundle adjustment, respectively. Section 6 describes the steps involved in the image formation, and Section 7 provides results using the real airborne data set. Conclusions are provided in Section 8.

2. AIRBORNE DATA SET

Figure 1 illustrates the standard image collection scene from the Global Hawk UAV. A scene consists of 140 total frames collected in a serpentine pattern of 14 scans of 10 frames per scan. It scans these 10 frames at 30 Hz, and requires 0.2 seconds of turnaround time before scanning in the reverse direction again at 30 Hz. The angular field of view (FOV) of a frame is 0.3 degrees; hence a 10-frame scan has an FOV of approximately 3 degrees. A single frame contains roughly 1000 by 1000 pixels; therefore, a scene converted to a mosaic at full resolution would contain roughly 10,000 by 14,000 pixels. The UAV is typically flown at approximately 16,800 meters above ground level and 45 degrees off-nadir, predominantly roll about the flight direction with only a few degrees of pitch. The nominal ground sample distance (GSD) was approximately 0.11 and 0.16 meters in the cross-range and range directions, respectively.
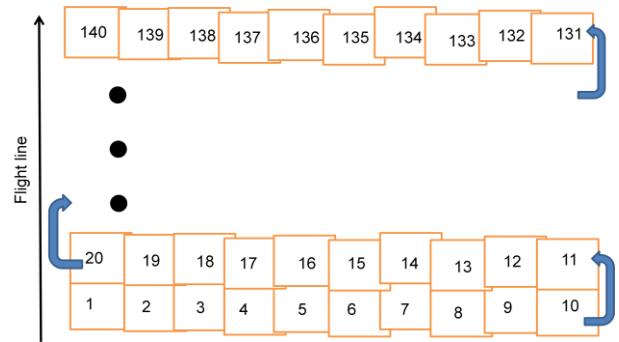

Figure 1. Image collection scene from the GH UAV

3. AUTOMATED TIE POINT MATCHING

The purpose of the automated tie point matching step is to find at least one tie point in each of the overlap regions between adjacent image chips. Imaging systems that scan to collect staggered arrays of chips, or use a whiskbroom collection to sweep an image chip, collect overlapping pixels that look very similar compared to an imaging system that collects a stereo pair with a wide convergence angle. In our case with scanners, the time difference between exposures for the pair of image chips is small enough that perspective differences, and therefore relative relief displacement effects, do not pose a challenge to the automated matcher. Likewise, scale, rotation, and radiometric differences are negligible. Furthermore, the metadata errors over the short time between subsequent image chip collections are highly correlated; while their absolute errors may be significant, the relative error between the two times is likely small. For all of these reasons an area-based matching technique, such as normalized cross correlation, is quite effective. Problems arise due to cloud cover or areas of non-distinct features such as would occur in the water, forest, or grassland. Automated tie point matching failures can be accommodated by using an outlier detection and removal technique, such as iterative re-weighting, in the bundle adjustment step. The automated matching process was not the main thrust of this research; instead, an industry partner provided the capability in the form of an executable file.

4. CAMERA CALIBRATION

Camera calibration is a critical step to ensure that features align along all borders of all chips. Performing a bundle adjustment with added parameters, each time an image is formed, is possible. However, we instead recommend performing the camera calibration once, and using the recovered interior orientation parameter values as constants in the bundle adjustment step. The next two sub-sections define the approach and parameter recovery, respectively.

4.1 Approach

The ideal configuration of airborne or spaceborne imagery to support camera calibration is to obtain a large number of overlapping images from geometries with large convergence angles and various rotations of the image chips about the optical axis. So it is a major advantage if the photogrammetrist has the ability to task the system that needs to be calibrated. However, typically imagery from these scanning systems is available only in its "mapping" mode that tries to efficiently sweep up as much ground coverage as possible by minimizing image chip overlap.

We did not have the ability to task the imagery collected from our airborne whiskbroom data set; therefore, we had to take full advantage of our image chip overlap areas. Specifically, we manually measured tie points not only in the overlap areas between adjacent chips but also measured all possible four-ray tie points. Thus we used many more and better tie points to perform the one-time calibration than we used in the bundle adjustment step that we would run for each image formation instance.

4.2 Recovery of Parameters

Satellite imagery camera calibration of the staggered array typically consists of a laboratory calibration, checked with real in-situ data after launch, to obtain precise field angle mappings for the ends of each image chip. However, for our airborne whiskbroom case, the camera calibration algorithm is a bundle adjustment with added parameters. Hence, the algorithm is the same as provided in Section 5 with the exception that a single set of up to 10 interior orientation parameters (IO) are recovered for the whole block of images. The following equations are used to calculate x and y corrections to the image points as a function of the IO parameters (Fraser, 1997).

$$\Delta x = \frac{\bar{x}}{c}\Delta c + \bar{x}r^2 K_1 + \bar{x}r^4 K_2 + \bar{x}r^6 K_3 + (2\bar{x}^2 + r^2)P_1 + 2P_2\bar{x}\bar{y} + b_1\bar{x} + b_2\bar{y}$$
$$\Delta y = \frac{\bar{y}}{c}\Delta c + \bar{y}r^2 K_1 + \bar{y}r^4 K_2 + \bar{y}r^6 K_3 + 2P_1\bar{x}\bar{y} + (2\bar{y}^2 + r^2)P_2$$
$$\bar{x} = x_b - x_o$$
$$\bar{y} = y_b - y_o$$
$$r^2 = \bar{x}^2 + \bar{y}^2$$

(1)

in which:

x_b, y_b are the observed image coordinates in the image coordinate system,

x_o, y_o, c the principal point offsets with respect to the image coordinate system, (x_o, y_o), and the camera principal distance,

Δc principal distance correction parameter,

K_1, K_2, K_3 three radial lens distortion parameters,

P_1, P_2 two decentering lens distortion parameters, and

b_1, b_2 two in-plane distortion parameters (scale differential and skew).

Results of a camera calibration for a real airborne whiskbroom imagery data set are provided in Section 7.2.

5. BUNDLE ADJUSTMENT

The purpose of the bundle (aka block) adjustment is to refine the values of the adjustable parameters, which are corrections to position and attitude of the sensor, associated with each image chip. The adjustment serves to globally minimize the sum of the squared weighted tie point image coordinate measurement residuals, thereby preparing the data for image re-sampling that will result in good alignment across the image chip boundaries. The next three sub-sections describe the algorithm, and its inputs and outputs, for the airborne whiskbroom case.

5.1 Inputs

Obvious inputs to the block adjustment, following from the discussion in the last two sections, are the image coordinates of tie point measurements and the IO values of the camera calibration.

Another required input is reasonable values for the exterior orientation (EO) parameters associated with each frame. Sometimes they are not provided explicitly in the metadata of the imagery and need to be estimated using geometric techniques. One example was in our specific data set in which the coordinates of the camera perspective centre were reliable while the camera attitude was not. Consequently, we used the latitude, longitude, and height coordinates associated with the four image corners to back out the three Euler angles describing the camera attitude. In another data set, not provided in this paper, the only metadata available with the images was cubic rational polynomial function coefficients which provide a means to perform image-to-ground and ground-to-image transformations. To solve for the perspective centre of the camera, we generated 3D lines in object space that represent imaging loci and then used least squares adjustment to find the best estimate of the intersection point where the imaging system was located during acquisition. The endpoints of lines defining the imaging loci were obtained by simply running the image-to-ground function at a nominal height and then again at a new height, and then generating a line through these points.

An optional input to the algorithm is a digital surface model (DSM); not having one available just means that ranges or terrain heights need to be estimated during the algorithm as discussed in the next sub-section.

Another optional input is error covariance data for all of the aforementioned inputs. Availability of this uncertainty information has many advantages. First, it means that the *a posteriori* error covariance matrices, which are an automatic by-product of the bundle adjustment, can be used to generate valid *a priori* error covariance matrices describing the idealized metadata associated with the output image product. Second, it provides correct relative weighting among all input measurements, thereby resulting in an optimal least squares adjustment solution. Finally, assigning reasonable uncertainties, instead of excessively large ones, to the adjustable parameters will prevent image warping and will usually make outlier detection procedures more effective.

5.2 Algorithm

The bundle adjustment follows closely to what can be found in photogrammetry text books (Mikhail, et al., 2001). The main difference between our adjustment and a standard one is that the pairs of imaging rays that determine a tie point have a very weak convergence angle. In order to eliminate instabilities in the solution, one out of the three coordinate components of the tie point must be constrained. The first choice, if a DSM is available, is to constrain the Z (height) component such that the tie point lies on the DSM. The second choice is to judiciously apply some *a priori* weight to one component of the tie point. Instead of choosing between the X (East), Y (North), and Z (Up) components, we chose to constrain the W component of tie points in the UVW coordinate system in which the W axis aligns with the nominal line-of-sight direction, the U axis aligns nominally with the vehicle velocity vector, and the V axis completes a right-handed coordinate system. In our real data, the sensor scanned cross-track so the time difference was approximately 1/30th second and ½ second between adjacent frames in the cross-track and along-track directions, respectively. Therefore, we allowed more movement (less weight) on the along-track tie points compared to the cross-track ones which were held essentially fixed in the W direction.

5.3 Outputs

The two main outputs of the bundle adjustment are refined values for the EO parameters of each frame, and a grid of interpolated slant ranges from the sensor to the ground. This grid of interpolated ranges is derived as a function of the recovered EO parameters and tie point ground coordinate values. The utility of the grid of ranges, aka range image, will become apparent in Section 6.3.

6. IMAGE FORMATION

The process of image formation requires first making a decision as to what the ideal, or virtual, image geometry will be for the desired output image product. Then, it requires assignment of ideal IO and EO parameters to the virtual image geometry. These steps, followed by the pixel resampling, are the topics of the next three sub-sections.

6.1 Defining the Virtual Geometry

As mentioned in the introduction, most vendors build mosaics by stitching the image chips together using 2D-to-2D transformations; thus, the formed image retains no mechanism to perform image-to-ground or ground-to-image mappings. A virtual geometry model, however, consists of parameters that allow the calculation of a perspective centre location and line-of-sight (LOS) vector associated with any pixel in the output image. The perspective centre location is calculated from the position components of EO, while the LOS vector is calculated from the attitude components of EO in addition to the IO parameters. Due to the nature of forming an output image using a virtual geometry, per Section 6.3, it is obvious that it perfectly replicates the original image geometry when the point of interest lies at the same elevation (or slant range from the sensor) that was assumed during image formation. As the error in elevation (or slant range) increases, the geolocation error increases. The sensitivity of geolocation errors to elevation (or slant range) errors is a function of how closely the virtual imaging geometry matches the actual imaging geometry.

6.1.1 Spaceborne Synthetic Linear Array Geometry: The choice of virtual geometry to use for the output image can be obvious, e.g. for a spaceborne staggered array scanner, or can require good engineering judgment as in the case for the GH UAV real data set described in Section 2. In the spaceborne staggered array case, the virtual image geometry is a linear array that lays half way between the leading and trailing rows of detectors that comprise the staggered arrays; see Figure 2.
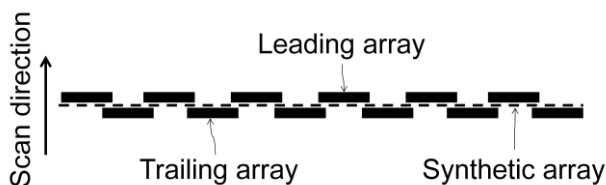

Figure 2. Staggered and synthetic arrays for a spaceborne system

6.1.2 Virtual Geometries from Airborne Frame: Using the 140 frames of the GH UAV as an example, the two extreme cases to output an image product from a ground station are to provide all of the original frame images, or to generate a single ortho-mosaic from all of the frames. The former case would allow for the most rigorous exploitation of the imagery, but the least convenient. The latter case, however, would result in the least rigorous exploitation of the imagery but would be the most convenient product to someone who wants it inherently co-registered with other map-projected products and does not want to use a sensor model. The remainder of this sub-section discusses three example virtual image geometry options; and the preferred method for this airborne data set, i.e. whiskbroom, was used to generate the results in Section 7.

6.1.2.1   Panoramic Geometry: The most rigorous virtual image geometry model would be to model the scene as 14 panoramic scans; therefore, the ground station would build 14 output image products, instead of 140 original frame images. Figure 3 illustrates the panoramic geometry for a single 10-image scan. Recall, from Section 1, that the GH UAV captures these 10 frames in one-third of a second. So, if the vehicle is travelling at 60 meters/second, then it has only travelled 20 meters between the first and tenth frames, i.e. plus-or-minus 10 meters worst case deviation from the virtual instantaneous perspective centre.
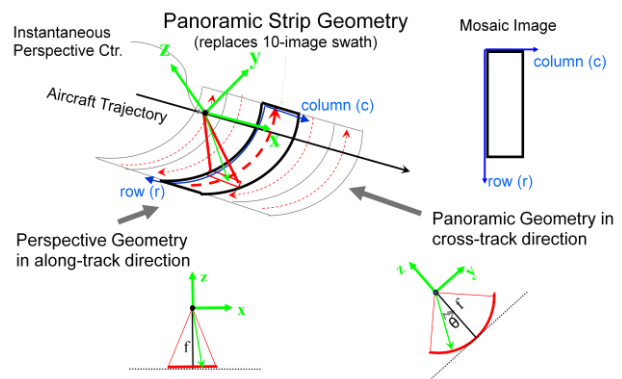

Figure 3. Panoramic geometry applied to 10-image scan from a GH UAV

6.1.2.2   Whiskbroom Geometry: The whiskbroom geometry is the most rigorous virtual model that can be applied to the entire scene, as a single output product. A single output product with a single sensor model is much more convenient than 14 output products with 14 separate instantiations of a panoramic model. If the 14 output products were concatenated into a single product, then the far end of the FOV would have overlaps while the near end would have gaps; see the left side of Figure 4. In the whiskbroom geometry model, all pixels in a given column of the output image correspond to a particular sensor position and sensor attitude; see Figure 5. Consequently, the whiskbroom model removes the effect of unequal scales in the near and far ends of the FOV, unavoidable with a panoramic model, by gradually compensating for the difference on an unnoticeable pixel by pixel basis; see the right side of Figure 4.
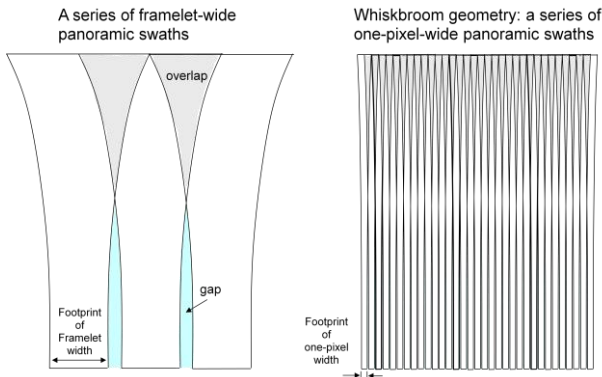
Figure 4. An exaggerated view of footprints of three panoramic images (left side) and several columns of a whiskbroom image (right side)
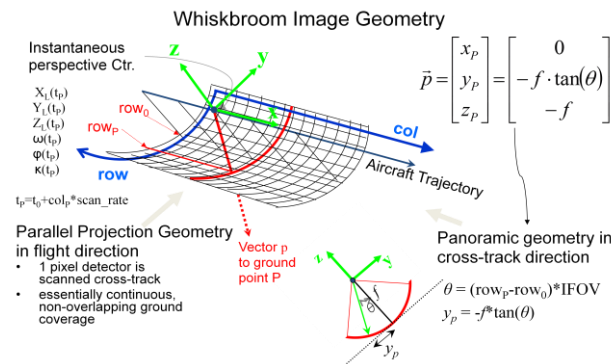

Figure 5. Whiskbroom geometry applied to a GH UAV scene

6.1.2.3 <u>Frame Geometry</u>: The frame geometry is a simpler but less rigorous virtual model that can be applied to the entire scene. It was deemed inappropriate for the GH UAV real data set due to the relatively long distance travelled between the first and last images collected in a scene. The virtual frame model may be appropriate from a platform at longer slant ranges, for smaller scenes, or for faster collections. The virtual frame model is appropriate for systems such as wide area airborne surveillance (WAAS), whereby all frame images are acquired at the same time from the same platform.

6.2 Generating the Idealized Metadata

Section 6.1 presented three different options for the virtual image geometry; and the whiskbroom geometry was chosen for the GH UAV data. All pixels in a given column are modelled as though they were imaged at the same time instant; hence our desire is to have a polynomial function that yields each EO parameter as a function of column number. Figure 6 shows the values of the attitude parameters recovered from the bundle adjustment (aka triangulation), plotted as the blue curves. The red curve illustrates the result of fitting a second order polynomial to the average value for the scan, i.e. the midpoint between the value at the fifth and sixth frames. The same polynomial fitting strategy was applied to the perspective centre positions. Unlike the case for the attitude parameters, these position parameters remained smooth throughout the duration of each scan.
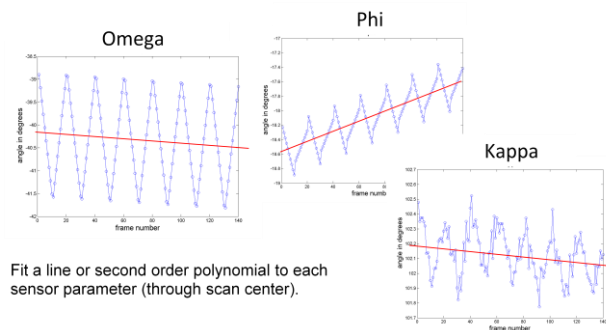

Figure 6. Smoothed attitude data

It is important to note that smoothing the EO parameter data can have significant advantages, as well as possible drawbacks, when employed during the image formation process. The following comments about smoothing apply equally to airborne framing, pushbroom/whiskbroom, or spaceborne linear array scanners. If the sensor had experienced uncommanded (but recorded by the IMU, or recovered during bundle adjustment) rolling during the image collection, e.g. due to air turbulence, then straight lines would appear as wavy curves in an image formed using this erratically varying attitude; however, smoothing the metadata of the virtual (idealized) image would have the effect of straightening the lines in the formed image. Similarly, any imperfections in sensor IO, e.g. modelled lens distortions or chip misalignments, need not be included in the virtual (idealized) imaging geometry since it incorporates unnecessary complication into the virtual geometry model which will ultimately need to be implemented by a downstream exploitation tool. Even substantial smoothing of the attitude data, or simplification of the IO parameters, will retain the geometric integrity of the output image, albeit with changes to how scale varies in the output image compared to the original image chips. (The reason why geometric integrity is retained, as explained in Section 6.3, is that the re-sampling process requires the unsmoothed attitude and imperfect IO parameters when performing the ground-to-image step.) Caution must be exercised, however, when smoothing perspective centre locations since this simplification will introduce errors into subsequent geolocation processes as a function of the amount of error in the DSM (or range image) that was used in the re-sampling process. Finally, note that while smoothing the attitude data of the virtual image geometry will have the desirable effect of making object straight lines appear as straight lines in the image (even though they appear as wavy curves in the original image), it will result in the artefact that the straight edges marking the extent of the original images will appear as wavy edges in the output image product.

6.3 Re-sampling the Pixels

Once the geometry of the virtual (idealized) image has been defined and the associated IO and EO parameters have been determined, the remaining step of re-sampling the pixels is relatively straightforward. First, it is recommended to add a buffer of black pixels around all four edges of the output image product. (It is a simple procedure to crop the image if a customer does not want to see the wavy or jagged edges in the final output image.) For each pixel (row, column) in the output image, the following steps are performed:

1. <u>Image-to-ground in the virtual image.</u> Calculate time as a function of row and column, obtain the values of all EO parameters as a function of time (from the

potentially smoothed functions), and then project the ray to the ground using either the DSM (if available as an input to the algorithm) or the interpolated range image (obtained as a by-product of the bundle adjustment) to provide the third dimension.

2. Ground-to-image (g2i) in the original image chip. Perform the g2i function into the original image chip using the detailed IO parameters obtained from camera calibration, and the EO parameters (unsmoothed) obtained from the bundle adjustment. If the point exists in multiple original image chips, then choose the one whose image coordinates lie farthest from the edge. If the point does not exist in any of the original image chips, then in the next step assign the pixel intensity as black.

3. Obtain the pixel intensity to place in the output image product row and column location. Use bi-linear interpolation from the four nearest pixels in the original image chip to obtain an intensity value. Nearest-neighbour or other interpolation techniques may be chosen, instead of bi-linear, to meet the desires of the user.

7. RESULTS

7.1 Camera Calibration

Section 4 described how camera calibration was performed on the GH UAV dataset. Figure 7 shows the image coordinate residuals after running the self calibration with no IO parameters (left side), compared to the residuals after running self calibration with focal length and de-centering lens distortions (right side). The blue arrows are the image coordinate residuals, while the red arrows are simply the signed mean of the cluster of blue arrows. The blue arrows consistently fall in four quadrants of a frame due to the design of the automated tie point finder mentioned in Section 3. The red arrows were intentionally shifted away from the blue arrows to improve readability. Note that camera calibration had the effect of reducing the systematic error from 2.4 to 0.1 pixels.
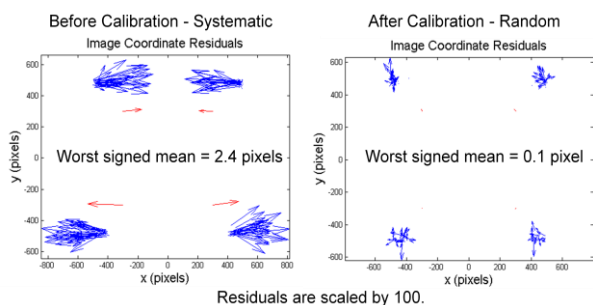

Figure 7. Image coordinate residuals at the 4 corners of each frame, before calibration (left) and after calibration (right)

Section 7.2 provides a comparison of output images produced using a calibrated versus uncalibrated camera.

7.2 Image Formation

While the algorithm documented in this paper was run on several data sets, this section provides two image product examples; see Figures 8 and 9. Since the figures had to be reduced in resolution to fit them in the paper, Figure 10 provides a zoomed in view of the Figure 9 image.

Image: AIS24H52_001118
Input images provided by Global Hawk Program Office
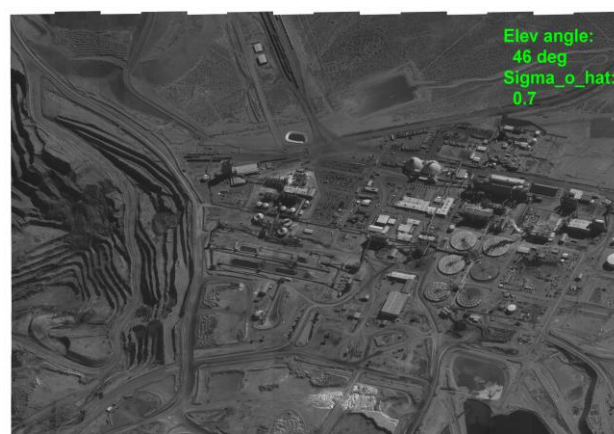Figure 8. GH UAV output image product, Example 1

Image: AIS24H83_002178
Input images provided by Global Hawk Program Office
Figure 9. GH UAV output image product, Example 2

Figure 10. GH UAV - zoomed in version of Example 2. Notice undesirable shift in linear features in Uncalibrated scene.

7.3 Registration

As discussed at the end of Section 1.2, exploitation of the final output image product is performed by implementing the same virtual image geometry model, i.e. whiskbroom, which was used to form the image. The output image can be registered to a reference base, i.e. a controlled stereopair, by incorporating adjustable parameters into the virtual whiskbroom model and

performing a photogrammetric resection using points extracted from the stereopair as ground control points (GCPs).

The adjustable parameters consisted of one interior orientation parameter, i.e. focal length, and coefficients of a polynomial function of time used to compute small corrections to the roll, pitch, and yaw angles about the x, y, and z sensor coordinate system axes, respectively, i.e. the green axes in Figure 5. The polynomial functions are:

$$\Delta\omega = a_0 + a_1 t + a_2 t^2$$
$$\Delta\varphi = b_0 + b_1 t + b_2 t^2 \qquad (2)$$
$$\Delta\kappa = c_0$$

in which:

$\Delta\omega, \Delta\varphi, \Delta\kappa$ are the roll, pitch, and yaw angles, respectively,

a_0, b_0, c_0 are constant terms of the polynomials,

a_1, b_1 are first order terms of the polynomials,

a_2, b_2 are second order terms of the polynomials, and

t is the normalized time associated with the column number (sample) of the coordinate of interest in the formed image; thus t ranges smoothly from -0.5 to +0.5 from the first to the last column of the image.

One hundred and fifty conjugate points were manually measured on the output image and on a WorldView1 stereopair that had a GSD of 0.5 meters. The output image had roughly the same GSD as the input image frames whose geometry was described in Section 2. Figure 11 shows the image coordinate residuals when 150 points are used to control the resection and the same 150 points are evaluated as check points. It refers to the baseline case of six adjustable parameters, whereby the second order terms in equation 2 were omitted from the adjustment. While 150 GCPs are not practical, the example was run in order to show what relative deformations exist in the output image. Note how closely that the mean theoretical ellipse matches the empirical ellipse. The former ellipse was derived by performing error covariance propagation on the single-image geolocation (to compute X and Y) for each ground point, using the GCP's Z value as input, and then averaging all of the horizontal error covariance matrices.
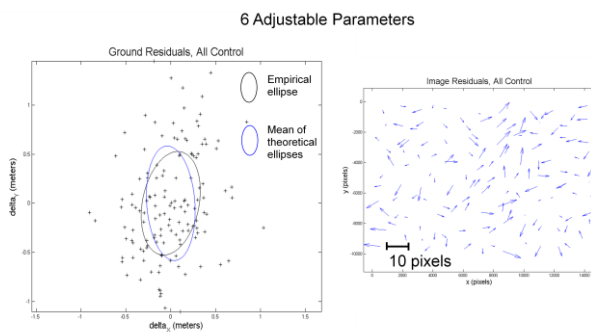

Figure 11. Ground and image coordinate residuals after resection with 6 adjustable parameters

Table 1 shows the values of the semi-major ("a") and semi-minor ("b") axes of the empirically calculated check point error ellipses for the 150 GCP case as well as practical cases with only 4 or 5 GCPs. In these latter cases, the check points consisted of only the points that were not used as GCPs. The reduction in number of control points, from 150 to a practical number of 4, did not cause the accuracy to worsen significantly,

thus indicating that good internal relative geometry has been preserved during the image formation process. The *a posteriori* reference variance values, provided in the last column of Table 1, are close to unity therefore indicating that the one-sigma uncertainties of 1 pixel and 0.5 meters for image coordinate measurement and GCP uncertainty, respectively, were appropriate. They also show that the adjustable parameter of second power for pitch is significant, but not roll.

Case	Semi-major, "a" (meters)	Semi-minor, "b" (meters)	Sig_o_hat
All pts	0.54	0.31	0.96
All pts*	0.50	0.26	0.82
All pts**	0.49	0.26	0.81
4 GCPs	0.57	0.52	1.34
5 GCPs	0.58	0.45	1.34
5 GCPs*	0.58	0.31	1.05

* Used 7 adjustable parameters (added second power for pitch)
** Used 8 adjustable parameters (added second power for roll/pitch)

Table 1. GH UAV – resection results

8. CONCLUSIONS

This paper provided the image formation concepts and steps to build products from airborne and spaceborne digital array scanners, i.e. automated tie point matching, bundle adjustment, defining the virtual (idealized) image geometry, and re-sampling the pixels. It emphasized the importance of performing camera calibration prior to using the resulting parameters to form image products. This photogrammetric approach to image formation was verified using real GH UAV frames to generate a mosaic product that could be rigorously exploited using its sensor model to perform processes such as resection, triangulation, and geolocation with error propagation.

While numerous vendors, e.g. commercial satellite optical image data providers, already know how to form images, this paper highlighted a few new concepts to some readers:

1) Although the sensor model associated with the raw image may be complex in that it contains many optical or focal plane distortions, the virtual geometry model can designed to be quite simple;

2) Although the time history of adjusted attitude data associated with the raw images may be rough, e.g. due to air turbulence, the metadata of the virtual geometry model can be designed with smooth variation; and

3) These virtual geometry models, and their associated metadata elements, can be generalized such that multiple vendors use the same models, therefore allowing users to exploit imagery from different vendors using the same generic sensor model.

9. REFERENCES

Fraser, C. S., 1997. Digital camera self-calibration, Invited Review Paper, *ISPRS Journal of Photogrammetry & Remote Sensing*, Vol. 52, pp. 149-159.

Mikhail, E., J. Bethel, J. McGlone, 2001. *Introduction to Modern Photogrammetry*. John Wiley and Sons, Inc., pp. 130-146.