A PROCEDURE FOR ORTHORECTIFICATION OF SUB-DECIMETER RESOLUTION IMAGERY OBTAINED WITH AN UNMANNED AERIAL VEHICLE (UAV)

Andrea S. Laliberte, Remote Sensing Scientist
Craig Winters, GIS Specialist
New Mexico State University
Jornada Experimental Range, Las Cruces, NM 88003
alaliber@nmsu.edu
craigwin@nmsu.edu

Albert Rango, Research Hydrologist
USDA-Agricultural Research Service
Jornada Experimental Range, Las Cruces, NM 88003
alrango@nmsu.edu

ABSTRACT

Digital aerial photography acquired with unmanned aerial vehicles (UAVs) has great value for resource management due to flexibility and relatively low cost for image acquisition. The very high resolution imagery (5 cm) allows for mapping bare soil and vegetation types, structure and patterns in great detail. While image acquisition is relatively straightforward, the creation of orthorectified, GIS-ready image mosaics presents multiple challenges. Those include relatively small image footprints, image distortion due to the use of low-cost digital cameras, difficulty in locating ground control points and in automatic generation of tie points, and relatively large errors in exterior orientation (camera position and attitude information from the UAV’s GPS/IMU). We developed an automated procedure to improve the accuracy of the exterior orientation by matching the UAV images to an orthorectified reference image. Using the UAV reported exterior orientation and camera geometry, combined with the reference image and DEM, the algorithm simulates image acquisition and then computes the covariance between camera image and simulated image pixels. With this evaluation function, a heuristic search algorithm finds successive improvements to the external orientation, ultimately producing a corrected exterior orientation that allows orthorectification with minimal input of tie points and/or ground control points. The RMS error for a 5-cm resolution, 257-image mosaic was 48 cm. Cost and turnaround time for production of orthorectified mosaics from UAV imagery are considerably reduced due to less time and money spent on ground control point and manual tie point collection.

INTRODUCTION

Sub-decimeter resolution aerial photography acquired with unmanned aerial vehicles (UAVs) has great potential for rangeland monitoring and assessment. Due to the relatively low flying height, spatial patterns and structure of vegetation and soil can be mapped in great detail (Laliberte et al., 2007; Rango et al., 2006), bridging the gap between ground-based rangeland monitoring protocols (Herrick et al., 2005) and remotely sensed information from aerial photos or satellite imagery (Laliberte et al., 2004). UAVs are less costly than piloted aircraft and can be deployed quickly and repeatedly, making them ideal for change detection at fine scales.

While image acquisition and image classification are relatively straightforward (Laliberte et al., 2007), multiple challenges exist for orthorectification of the individual images for the purpose of creating a GIS-ready mosaic. Due to the low payload capability of most small UAVs, imagery is often acquired with low-cost, off-the-shelf digital cameras. Imagery from those cameras has greater distortion compared to imagery from mapping cameras, and a camera calibration is required to determine the camera’s interior orientation parameters. If sufficient ground control points of good accuracy are available, a self-calibrating bundle adjustment can be performed (Wu et al., 2006). Otherwise, camera calibration parameters can be derived by taking images of a test field (Clarke and Fryer, 1998).

Although the UAV imagery has very high resolution (~5 cm pixel resolution at a flying height of 150 m above ground), the image footprint is relatively small (152 m x 114 m), and the images have to be stitched together to create a mosaic for further analysis and classification. The small footprint and very high resolution make it difficult
to correlate image features to coarser resolution reference imagery, such as digital orthoquads (DOQ). The 1 m resolution of the DOQ constrains the accuracy of the individual control points and also makes the matching of points between the reference image and the UAV image a visually intensive manual operation. In addition, our study area is located in desert shrubland with few distinguishing features (as are most rangelands), and this creates problems for the image analyst for finding manual tie points and ground control points (GCP) on overlapping imagery for photogrammetric processing. A direct georeferencing workflow (Jacobsen, 2002) would be desirable with this imagery, but the relatively large error of the UAV’s position and attitude information from the onboard GPS/IMU currently precludes this approach. In previous orthorectification projects with this type of imagery, the image center coordinates determined by the UAV’s GPS unit had an average 30-m along track error, due to lack of differential correction and to time lags between the camera trigger and the GPS (Laliberte et al., 2007). Because the error of the exterior orientation (EO) parameters are at least an order of magnitude larger than the error associated with manually derived tie points and GCPs, automatic tie point collection in photogrammetric software is prone to errors or fails entirely in many cases.

In previous attempts, we only achieved reasonable orthorectification results with this imagery (RMS error from aerotriangulation: 0.33 pixels) by omitting the EO altogether and relying on manual tie points and GCP collection. However, this is not a cost-effective solution due to the time it takes to collect those points manually. With hundreds of images per block, determining sufficient manual tie points and GCPs represents a significant human resource component in an imaging system that can collect thousands of images in a day of flying. Our objective was to develop a semi-automated approach for orthorectification in order to reduce the cost and turnaround time for production of orthorectified UAV mosaics. Specifically, we focused on improving the accuracy of the EO parameters, and minimizing the number of manual tie points and GCPs, so that the images could be processed in photogrammetric software to produce an orthorectified mosaic.

**UAV PLATFORM, SENSORS AND IMAGE ACQUISITION**

We used a MLB BAT 3 UAV (MLB Company, 2006)\(^1\), a small UAV with a 1.8 m wingspan and 10 kg weight. The BAT system consists of a fully autonomous GPS-guided UAV, a catapult launcher, ground station with mission planning and flight software, and telemetry system (Figure 1). The GPS module (TIM-LP Antaris\(^1\))\(^1\) has an update rate of 4 Hz, and the accuracy is 2.5 m CEP, but the data is not differentially corrected. The accuracy of the attitude data from the inertial measurement unit (IMU) is rated as ± 2 degrees for roll and pitch, and ± 5 degrees for heading. Two sensors are on-board: a color video camera with optical zoom capable in-flight and live video downlink to the ground station, and a Canon SD 550 7 megapixel digital camera. The BAT has an endurance of 2-6 hours, but the current limitation is the capacity of the camera’s flash card. We can store 1200 images on the camera’s 4 GB

---

\(^1\) Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the US Department of Agriculture.
flashing card, limiting the flying time to approximately 2 hours. The flight mission is planned on the ground station, and the UAV acquires imagery at 60-70% forward lap and 30-40% sidelap. The onboard computer records a timestamp, GPS location, elevation, pitch, roll, and heading for each acquired image, and this data is downloaded after landing. The inertial data stored for each image may differ from when the camera took the picture by up to 1 second, resulting in the observed discrepancies between actual and recorded image locations.

Imagery was acquired in October 2006 at a flying height of 150 m above ground over an arcuate-dune landscape at the Jornada Experimental Range in southern New Mexico. The camera’s resolution was 3072 x 2304 pixels with a field of view of 53.1 degrees, resulting in an image footprint of 152 m x 114 m with a pixel resolution of 5 cm. The UAV flew 8 parallel flight lines, acquiring 257 images and covering an area of 1800 m x 950 m. In order to determine the camera’s interior orientation (IO) parameters (radial lens distortion, principal point offset, focal length) (Fryer, 1996), we performed a camera calibration by taking photos of a calibration grid from various angles using the same camera settings as those used during aerial photo acquisition. PhotoModeler Pro 5 (Eos Systems Inc., 2006) was used to derive the camera’s IO parameters.

METHODS

Overview of the Method

The basic approach consists of combining the reference image (DOQ), a 10 m DEM, the initial EO, the camera’s IO parameters, and initial tie points into a sensor/environment model, and then to simulate image acquisition. Simulated imaging uses the initial EO recorded by the flight control system to extract a patch of the reference image and transform it to match the sensor array view. This simulated image is compared with the actual sensor image and evaluated for its match. The EO is repeatedly adjusted by various means and each new EO is likewise evaluated by producing a simulated image and evaluating its match to the sensor image. When the correction script has run its course, the best scoring EO is recorded as the corrected EO for the image. Those EO parameters are then used as input to the Leica Photogrammetric Suite 9.1 (LPS) (Leica Geosystems) software for orthorectification. The main steps of the EO improvement module we call PreSync are 1) initial tie point alignment, 2) rigid block adjustment, 3) independent registration of each image, and 4) realignment of tie points. A flowchart of the entire process is shown in Figure 2.

![Flowchart of UAV image processing for EO adjustment with PreSync and orthorectification with LPS.](image)

Figure 2. Flowchart of UAV image processing for EO adjustment with PreSync and orthorectification with LPS.
Tie Point Generation

The first step, the generation of automatic tie points, was done with the commercial photo panorama tool Autopano (Kolar Autopano, 2007). Feature matching in Autopano is based on the scale invariant feature transform (SIFT) algorithm developed by Lowe (2004). This algorithm does not require initial EO data or image-to-flight line matching, it is less sensitive to roll, pitch and heading changes from image to image than correlation or feature based matching algorithms, and it has been used successfully in photogrammetric applications (Laebe and Foerstner, 2006). The SIFT algorithm outperformed a number of other local descriptors in evaluations (Mikolajczyk and Schmid, 2005), although our experience has shown that both LPS and Autopano can produce incorrect tie points.

The PreSync Module

Search Space and Evaluation Metric. The search space for the correct EO is a six dimensional space represented by the EO vector: (X, Y, Z, ω, φ, κ). Our experience has shown that a reasonable search space, referred to as the global search space, extends roughly +/-50m spatially, +/-10° in ω and φ, and +/-20° in κ from the initial EO reported by the flight control system. As the search proceeds and the EO is approaching the correct value, a smaller space around the current EO is referred to as a local search space. The degree of match between the sensor image and the simulated image is determined by evaluating the covariance between the two image arrays, with a higher covariance corresponding to a greater match between the two images. Two methods have been developed to systematically adjust the EO.

In the gradient search method, the gradient is followed to the covariance maximum. This method proved to be less useful than anticipated due to many local maxima, because the evaluation function fEO is the covariance of two images, not a smooth mathematical function in the usual sense. The empirical sense of the function value (visualizing the subspace as a surface) can be visualized as a gentle hilltop with lots of bumps. Wherever the EO is located, the gradient search will take it to the top of the nearest bump, and in our imagery, there is a local maximum within a few meters of any given location. Therefore, gradient following is used at the end of the search to “fine-tune” the solution.

To overcome some weaknesses of following the gradient to the nearest local extreme, the Downhill Simplex Search method (Press et al., 1988) was employed. This method is capable of stepping over nearby local maxima, and it makes good progress when there is a distinct global gradient. After converging, it can be restarted, again with the large simplex, and will converge again on a higher maximum. This can be repeated until the improvement is arbitrarily small.

Initial Tie Point Alignment. In PreSync, each tie point pass is an iterative procedure in two steps. The first step is to assign a ground coordinate to each tie point as follows: first, given a tie point, for each tied image, the image coordinate of the tie point is projected onto the ground using the current image EO. The assigned ground coordinate is the average of the ground coordinates for the given tie point. Second, given an image, for each of its tie points, the tie point image coordinate is projected onto the ground and the error to the tie point’s assigned ground coordinate is determined. Next, the EO of the given image is adjusted to minimize the total RMS error of the image’s tie points. When these steps are complete, each image has a new EO and the next iteration assigns new ground coordinates to the tie points as the process repeats. When the corrections from an iteration are arbitrarily small, the iterations are complete.

Rigid Block Adjustment. After tie points are aligned, image registration with the reference image may be off, but to independently adjust the images again would undo the image to image alignment. Our solution is to perform a rigid adjustment of the entire block of images. To do this, a principle point and EO are created at the center of the block. This EO is adjusted using the downhill simplex search method. Adjustments to the block EO are propagated to the individual image EOs and the adjustment is scored as the average of the scores of the individual images, with the objective of maximizing the average covariance of the block. This method preserves the image alignment while improving the registration between the UAV images and the reference image.

Individual Image Registration. It is conceivable that the entire process for a block of images could be reduced to first aligning tie points, and second, performing the rigid block adjustments as the only method for registering images with the reference image. We had reasonable results with this approach in terms of quality and processing time, but with a minor problem that simple tie-point alignment allows relatively large tip-tilt variability between images — a hinge-like flexing in the overlap area. Therefore we implemented additional steps to independently reference each image to the reference image, using the gradient following search method. Initial tie point alignment and rigid block adjustment get each image very close to its actual EO, while the individual image registration lets each image seek its preferred orientation within the local space.
**EO Correction Statistics.** The covariance evaluation between a sensor image and its simulated image provides a measure of how well the current EO matches the reference image. When working with a block of images, it is important to determine how uniform the EO correction is from one image to another. Due to a number of factors affecting the initial EOs (wind direction and systematic errors), the EO corrections tend to vary considerably more between flight lines compared to within a flight line. Therefore, for each flight line, we calculated the mean and standard deviation of the EO corrections for the images relative to the initial EOs. For each image, the deviation from the mean correction of its flight line, normalized by the standard deviation, yields a vector $EO_{corr}$, showing how far the correction of the image varies from the flight line’s mean correction in terms of standard deviations:

$$EO_{corr} = \left( \frac{\Delta X - \bar{\Delta X}}{\sigma_{\Delta X}}, \frac{\Delta Y - \bar{\Delta Y}}{\sigma_{\Delta Y}}, \frac{\Delta Z - \bar{\Delta Z}}{\sigma_{\Delta Z}}, \frac{\Delta \omega - \bar{\Delta \omega}}{\sigma_{\Delta \omega}}, \frac{\Delta \varphi - \bar{\Delta \varphi}}{\sigma_{\Delta \varphi}}, \frac{\Delta \kappa - \bar{\Delta \kappa}}{\sigma_{\Delta \kappa}} \right),$$

where $\Delta X$ is the EO correction in X for the image and $\Delta X$ and $\sigma_{\Delta X}$ are the mean and standard deviation for the image’s flight line (and similarly for the other EO elements). The magnitude of the $EO_{corr}$ vector provides a general detection mechanism for outliers. After the individual image registration is complete, images with an $EO_{corr}$ magnitude exceeding a given limit, typically 5.0, have their EO reset to the initial EO plus the mean correction for the flight line.

**Realignment of Tie Points.** The final step realigns the image tie points while maintaining the tip-tilt ($\omega$ and $\varphi$) corrections from the prior step. The tip-tilt corrections are not explicitly maintained, but within the angular range involved here, tip-tilt have less effect on tie point alignment than the other EO elements, and therefore small adjustments remain relatively undisturbed by this step. At this point, we perform another rigid block adjustment and, if further alignment is needed, we repeat the individual image registration and realignment of tie points to tighten up the alignment. On completion, the corrected EOs and tie points with associated ground coordinates are exported in LPS format.

**Implementation.** PreSync was developed using the Python language system (http://python.org/), including the Python Imaging Library (PIL) (http://www.pythonware.com/products/pil/) and the numerical python library (NumPy) (http://numpy.scipy.org/). This choice allows rapid flexible development needed for experimental work but has reasonable performance on computer intensive operations via the C-coded libraries.

**LPS Processing**

After importing the corrected EOs and tie points with ground coordinates into LPS, additional auto tie points can be generated to improve the aerotriangulation results. After orthorectification, the images are combined into a seamless mosaic.

An optional last step is the resampling of the mosaic with the AutoSync module in Erdas Imagine 9.1 (Leica Geosystems), using the 1-m resolution DOQ as a reference image. This step improves the alignment of the mosaic with existing imagery with little additional time involved, because the GCP collection in AutoSync is automated. The improved positional accuracy of the mosaic helps in overlaying field plots for relating ground measurements with remotely sensed information.

**Accuracy Assessment**

We assessed the geometric accuracy of the orthorectified mosaic by determining the coordinates of 72 independent check points of visible features with a Trimble Pro XR® differential GPS unit. The estimated accuracies for the differentially corrected positions for the checkpoints was 15% of the positions in the 15-30 cm range, 84% in the 30-50 cm range, and 0.4% in the 50-100 cm range. We then calculated the root mean square error (RMSE) between the 72 checkpoints and image points for two mosaic outputs: 1) the mosaic derived from the PreSync/LPS process (PreSync output), and 2) the resampled mosaic from the AutoSync module (AutoSync output).

**RESULTS AND DISCUSSION**

**Processing Time**

The resulting orthorectified mosaic is shown in Figure 3. The entire processing time was approximately 11 hours, broken down as 45 min for the Autopano tie point collection, 8 hrs for the PreSync run, 1 hr for orthorectification, 1 hr for mosaicking, and 30 min for the AutoSync resampling. Most of this time is hands-off, with total operator interaction estimated at 2 hours. The bulk of the processing time is the PreSync run, which
currently averages 2 minutes per image, but we are still in the process of optimizing the code, and are expecting to reduce the time of this processing step.

Figure 3. Orthorectified mosaic of 257 UAV images at the Jornada Experimental Range in southern New Mexico (left) with enlarged portion of red rectangle (top right), and comparison with 1 m resolution digital orthoquad image (bottom right).

Accuracy Assessment

The results of the geometric accuracy assessment of the mosaic are promising. Total RMS errors were 1.64 m before resampling the mosaic using AutoSync, and 0.48 m after that step (Table 1). We were limited in our error assessment by the accuracies of the differential correction process, which showed that most of the positions were in the 30-50 cm range, while trying to assess an image with 5 cm resolution. Future accuracy assessment will be conducted with survey-grade GPS to get within the range of 1 pixel error for the differential correction process.

Table 1. RMS errors for geometric accuracy assessment of mosaics from PreSync/LPS output and AutoSync output

<table>
<thead>
<tr>
<th></th>
<th>RMSE X</th>
<th>RMSE Y</th>
<th>RMSE Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>PreSync output</td>
<td>1.196</td>
<td>1.115</td>
<td>1.635</td>
</tr>
<tr>
<td>AutoSync output</td>
<td>0.372</td>
<td>0.301</td>
<td>0.479</td>
</tr>
</tbody>
</table>

The residual plot (Figure 4) shows the magnitude and character of the errors for both mosaic outputs. The PreSync/LPS output had a systematic displacement. Resampling this mosaic again with AutoSync tightened up the error and removed the systematic error in a simple step. Visual assessment of the results also included overlaying the UAV mosaics on the DOQ and assessing the alignment of image features, which proved to be a good visual fit. While the AutoSync resampling step tightens up the overlay of the mosaic with existing imagery, it has to be used with caution, especially in areas with larger changes in elevation. The PreSync/LPS output consists of mosaicked orthorectified imagery, while the AutoSync output resamples the entire mosaic as one image, even though a DEM is taken into account. In our case, the imagery was acquired over a relatively flat area, with a change in elevation of only 14 m, and we were not overly concerned with distortions associated with the last resampling step. For rangeland research, it is important that vector files of field plots can be overlaid on the mosaic with sufficient

ASPRS 2008 Annual Conference
Portland, OR • April 28 – May 2, 2008
accuracy for comparing field measures with remotely sensed information, and we feel confident that the accuracies achieved with our approach are sufficient for rangeland monitoring and assessment purposes.

Figure 4. Residual plot showing absolute differences in meters between coordinates of 72 check points measured with differentially corrected GPS and from image locations. Red points are for the mosaic created from the PreSync/LPS output, blue points are for the mosaic resampled additionally with AutoSync.

The Use of Tie Points
Initially, PreSync was run without the use of tie points, and while most images could find their locations relative to the reference image, a small percentage tended to “wander off”. Among the problems encountered were lack of detail within the frame area, regular feature repetition, producing large scale maxima at multiple locations, and changes in features such as vegetation cover, dunes or gullies between the acquisition time of the reference image and the UAV image. The search algorithms responded in various undesirable ways to these problems, including matching to a wrong ground location, decreasing the altitude to match the image to a small area with the average tonality of the reference image, or increasing $\omega$ or $\phi$ (roll/pitch) to the point that the image is matched with the no-data area beyond the reference image. To combat these problems and to achieve precision image matching for the final mosaic, the inclusion of image tie points was added to the process. We also investigated the use of LPS for auto tie point generation instead of tie points from Autopano. If LPS tie points were used, a preliminary PreSync run had to be performed to correct the initial EOs, because LPS can only generate auto tie points without manual intervention if the images have reasonably good EO values. A second PreSync run further refined the EOs using the LPS tie points. The advantage of using LPS is that with sufficient overlap, tie points can be found on more than two images, whereas Autopano produces exclusively two-image tie points. Because more estimates of the ground coordinate are averaged with LPS tie points, error is reduced. The disadvantage of this process is the additional PreSync step.

The accuracy of the tie points remains an issue, however, as it is limited by the accuracy of registration between the 5-cm resolution UAV images and the 1-m resolution reference image. Our current process involves using the tie points with associated ground coordinates exported from PreSync. Geometric accuracy can be improved by using the PreSync output as tie points only and by including an adequate number of accurately surveyed GCPs for aerotriangulation.

Search Space Methods
Our experiences with the gradient following, and downhill simplex search methods suggest other potential approaches. One option would be to run a low-pass filter over the image prior to the search to remove local maxima but maintain the global maximum. Another approach would be to initially generate a coarse grid of points spanning...
the global search space and to follow the gradient at each grid point producing an evenly distributed set of local maxima. Then, a quadratic equation can be fit to those values, and the maximum of that equation can be determined analytically or numerically. In theory, the solution will be a point near the evaluation function maximum.

CONCLUSION AND FUTURE RESEARCH

Our approach of improving the EO of UAV images for further photogrammetric processing has proven successful for a number of reasons. The process does not require much operator interaction; error assessment shows that the RMS value for the final mosaic is sufficient for rangeland research and vegetation classification; and cost and turnaround time for production of orthorectified mosaics are considerably reduced compared to traditional approaches, because less time and money is spent on GCP and tie point collection. Nevertheless, there is room for improvement and need for further research.

For this study, the imagery was acquired over a relatively flat area. Further test of the image processing approach in areas of higher relief are needed. In addition, a finer resolution DEM is needed for this processing approach, and we will investigate the potential of digital terrain model extraction from the imagery itself and its accuracy.

We have upgraded the digital camera to a Canon SD 900 with 10 megapixels, and we are currently acquiring imagery at 75% forward lap and 40% sidelap, which improves automatic tie point collection and ensures sufficient overlap in case of sudden wind gusts affecting the UAV. We are also investigating matching the UAV images to a QuickBird satellite image instead of a DOQ. Although there is an improvement in resolution with QuickBird imagery, the cost involved in acquiring the satellite images may be prohibitive in some cases, while DOQs are freely available. Therefore the QuickBird solution may not be feasible in all situations.

We are also in the process of comparing the described approach with a more traditional photogrammetric technique by acquiring images of a test field with targets located with survey-grade GPS. This will allow us to assess the difference in accuracy between the two approaches as well as the difference in camera interior orientation parameters derived from the external camera calibration compared to a self-calibrating bundle adjustment using the targets. Ultimately, our goal is to develop a workflow consisting of acquisition, orthorectification, mosaicking, and classification of UAV imagery suitable for rangeland monitoring at relatively low costs with quick turnaround times.

REFERENCES


