Photogrammetry Theory Companion
To be used with Photoscan Tutorial

1. Photogrammetry is the task of creating 3-dimensional maps using a series of overlapping 2-dimensional aerial images. In the graphic below, we can compute the 3D coordinates of the top of the tree (red dot) with 2 overlapping aerial images taken from exposure station 1, and exposure station 2. As you can see, the treetop is visible in both images but is seen from different perspectives. This perspective difference, aka parallax, enables us to construct a triangle between the treetop and the two exposure stations. We can then use geometry and trigonometry to calculate the dimensions of the triangle with the goal of estimating the x, y, and z coordinates of the treetop feature.

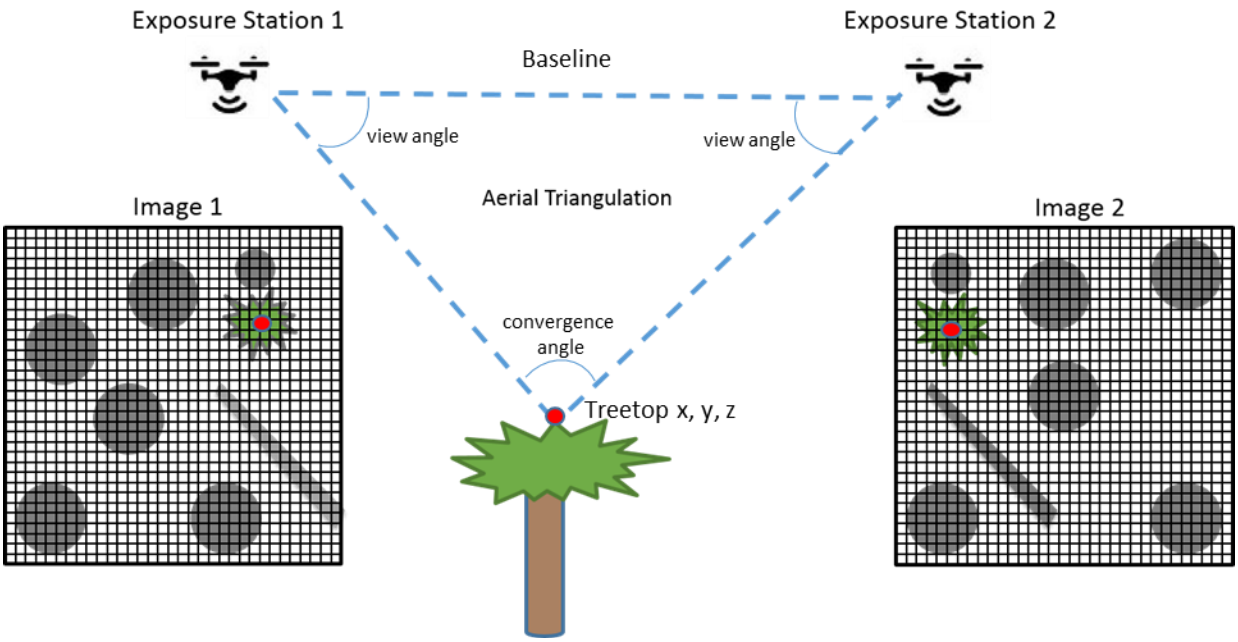


Fig. 1. Basics of aerial triangulation

2. Initial Photo Alignment

In SfM photogrammetry, an important aspect is to have **many** perspectives of the ground feature of interest, not just two. We call this multi-view stereopsis. By having many perspectives, we can create hundreds or even thousands of triangles for each ground feature. The first step in the workflow is to automatically identify thousands of ground features that can be seen in multiple images. Photoscan and other SfM software programs use scale invariant feature transform (SIFT) or something similar to match features across images with different scale, rotation, illumination and viewpoints. Next step is estimating the view angles of the triangle. These angles can be determined from identifying the specific pixel and therefore specific photodiode on the sensor array. For example, in Image 1 of Fig. 1, the treetop feature appears in column 26, row 10. Because light is assumed to travel in a straight line, we can estimate the view angle if we know other internal characteristics of the camera (focal length, sensor array size, lens distortion parameters).

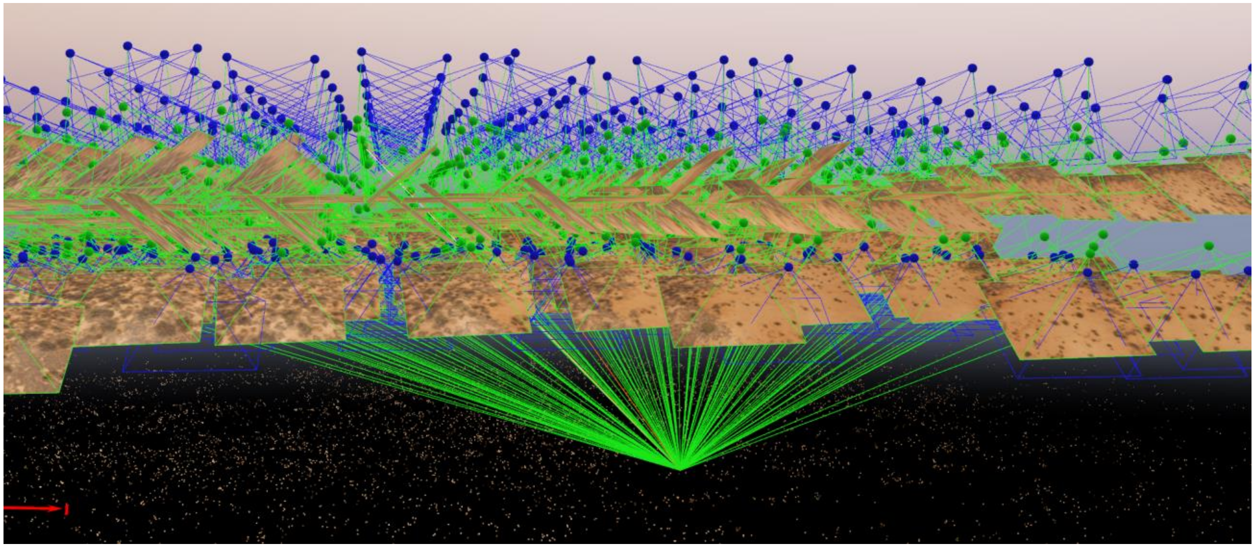


Fig. 2. One ground feature is found in dozens of overlapping images. Each ray (green line) connects the ground feature to the location of the feature in each aerial image. Image is screen shot from Pix4D software.

The procedure of matching ground points on many images and building triangles repeats thousands to millions of times until the relative dimensions of the scene take shape. This is done through an iterative least-squares process known as a bundle adjustment. Many images, perspectives (nadir & oblique), and even scales will strengthen the geometry of scene.
The outputs of the initial alignment are:

- Sparse point cloud with arbitrary relative coordinates (also called tie-points)
- Estimation of the pose of each camera (i.e., attitude)
- Self-calibration lens distortion parameters (need section on camera calibration)

The initial alignment and sparse point cloud generation using Sfm software is accomplished with no reference information (e.g., ground control). This is the magic sauce of SfM and a major departure from traditionally photogrammetry methods that rely on reference information to calculate initial alignment and camera pose. SfM methods are very flexible in terms of accepting images with inconsistent overlap, rotation between successive images, and images from different angles (i.e. nadir and oblique). SfM can also handle images collected with consumer grade cameras with unknown or unstable lens characteristics. Traditional photogrammetry is less flexible in that it requires consistent image overlap, minimal rotation between successive images, and calibrated sensors

3. As the algorithms are looking for ground features, you can change the spatial resolution or size of the ground features it is looking for. 'High' accuracy means it will try to find ground features at the resolution of the imagery. For example, if the resolution of the imagery is 2 cm, it will try to find and match ground features at that resolution. For lesser accuracies, the algorithm will coarsen the imagery before looking for ground features to match. For 'medium' accuracy, it resamples 4 pixels together into a single pixel before the matching routine begins.

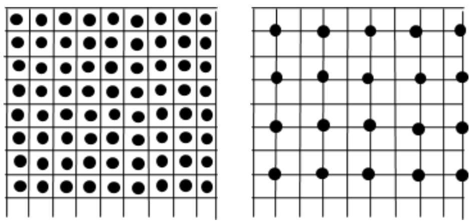


Fig. 3. High accuracy (left) and medium accuracy (right)

4. Generic preselection will speed up initial alignment by finding overlapping features and images by first doing it with a lower accuracy setting. Reference preselection uses the GPS coordinates to determine which images are likely to be overlapping and which are not, thus speeding up the process. For the reference preselection procedure to work effectively for oblique images, you must have the roll, pitch, yaw data for each image. This data is available in the EXIF file of the Phantom images, but I would recommend NOT pursuing it. It would be a hassle to mine the data out of the EXIF and get in into Photoscan. Please keep in mind that preselection procedures are meant to speed-up processing only, they should not affect the quality of the initial alignment.

5.

Key point limit: The number indicates upper limit of feature points on every image to be taken into account during current processing stage. Using zero value allows PhotoScan to find as many key points as possible, but it may result in a big number of less reliable points.

Tie point limit: The number indicates upper limit of matching points for every image. Using zero value doesn't apply any tie point filtering.

In my humble opinion, I would prefer to find a large amount of candidate tie points during the initial alignment. Some will be good but many will be poor quality. The poor quality points will be removed during the gradual selection process.

6. Why are GPS coordinates of image locations less accurate than ground control for referencing point clouds?

- The GPS equipment in the aircraft is often cheaper and less accurate compared with ground survey equipment

- The aircraft is a moving target and is only at the exposure station for a split second. On the ground, we can hold at a survey location for a long as it takes to get a great GPS signal.
- High precision ground GPS systems often use a base station and rover set-up to improve the accuracy. This same principle could be applied to drones. Currently, differential GPS is not available for Phantoms.
- There is an offset between the location of the GPS unit and the camera lens
- Synchronizing the camera exposure time with the GPS time can be challenging

7. Use a network of GCPs if you need the imagery products to be correctly georeferenced, scaled, and oriented. For example, if I am making a DEM for hydrological modeling, I would need to use GCPs to make the slope of the DEM correct.

If the goal of the imagery products is to estimate vegetation cover (%), then a scale-bar may be the most efficient way to reference the scene. To estimate cover, scale is important, but correct slope and georeferencing may be less important.

If I want to make point clouds, DEMs, or orthomosaics for mainly visual purposes (not measuring things) then using the image GPS coordinates could be a sufficient reference source. This approach may be useful if you want to ballpark map a small wildfire. I could also be useful to visually show phenology differences between with imagery sets taken during different seasons.

8. Setting the accuracies of the reference measurements are an important aspect of the least-squares bundle adjustment. The iterative process takes all of the input information including tie point locations, view angles, internal camera parameters, and reference information to create a best fitting 3D model. Reference accuracy tells the iterative process how far those measurements are allowed to adjust.

9. Optimization is a bundle adjustment using the reference information to improve the sparse point cloud and camera poses. This is more akin to traditional aerial triangulation methods. If the reference information is highly accurate, it can improve upon the sparse point cloud and camera poses estimated with the pure computer vision (structure-from-motion) approach.

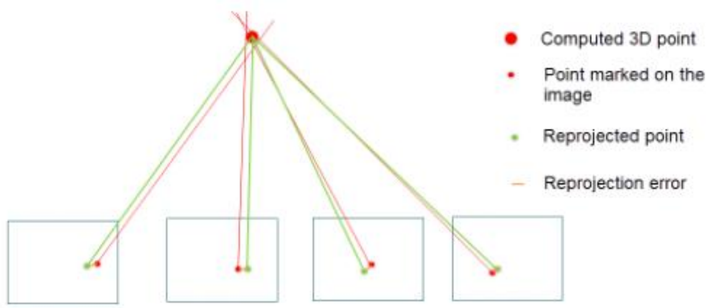
10. James et al. (2017) found that using focal length, principal point, 3 parameters of radial distortion, and 2 parameters of tangential distortion produced sparse point clouds with the lowest RMSE errors. Trying to optimize all camera parameters caused an overfitting that reduced model quality.

11. An electronic rolling shutter exposes the photo diodes line-by-line instead at the same time as found in a global shutter. If a camera is moving fast during the rolling shutter read-out, it can cause geometry distortion in the images. The Phantom 3 and 4 cameras have a rolling shutter read-out of 0.0033 seconds. Here is some additional information on rolling shutters.

12. Reconstruction uncertainty is based the shape of the triangles used to estimate the coordinates of the ground feature. You could also call this poor geometry. There is an ideal convergence angle for estimating highly accurate 3D coordinates. Good convergence angles should be 30° to 60°.

13. Projection accuracy is the pixel error that occurs when matching features across images. Very clear, crisp images could lead to better projection error, while blurry images could lead to higher projection accuracy.

14. Reprojection Error is the distance between the point on the image where a reconstructed 3D point can be projected and the original projection of that 3D point detected on the photo and used as a basis for the 3D point reconstruction procedure. High reprojection error usually indicates poor localization accuracy of the corresponding point projections at the point matching step. It is also typical for false matches. Removing such points can improve accuracy of the subsequent optimization step. Reprojection error is referring to 2D image coordinates in pixels. I believe that tie point accuracy is the same thing as reprojection error.



James, M. R., Robson, S., D'Oleire-Oltmanns, S., & Niethammer, U. (2017). Optimising UAV topographic surveys processed with structure-from-motion: Ground control quality, quantity and bundle adjustment. *Geomorphology*, *280*, 51–66. https://doi.org/10.1016/j.geomorph.2016.11.021