

PHOTOGRAMMETRIC 3D MEASUREMENTS IN FORESTS WITH ACTION CAMERAS

A. Pinhal^{1*}, J. A. Gonçalves^{1,2}

¹ University of Porto, Science Faculty, Rua Campo Alegre, 4169-007 Porto, Portugal – (apinhal, jagoncal)@fc.up.pt

² CIIMAR - Terminal de Cruzeiros de Leixões, 4450-208 Matosinhos, Portugal

Commission II, WG 4

KEY WORDS: Action camera, structure from motion, point cloud, positional accuracy, forest inventory.

ABSTRACT:

This paper presents a study of the application of action cameras, such as the GoPro, for the extraction of 3D models in a forest environment. These cameras are very small and can be used by any operator, even in other forestry works, to collect imagery for photogrammetric post processing. The method implemented uses the Structure from Motion approach in order to obtain the 3D model. The model becomes georeferenced by means of the camera positions which are recorded by the camera navigation GPS receiver. The aim of the study is to assess the geometric accuracy of such methodology, especially in relative terms, in the evaluation of object dimensions. Several tests were carried out both in open sky and under trees, with worse conditions for GPS positioning. The results pointed to scale accuracy of around 3%, which may be acceptable for many applications. Accuracy can be improved by using a scale bar of known length placed in the scene. The method is intended to be used in tree measurements for forest inventory or other outdoor measuring.

1. INTRODUCTION

Photogrammetry provides methods for three-dimensional measurement in a large diversity of conditions, both from aerial or terrestrial images. Its application in the forest environment is mainly done from aerial images, but there has been a recent interest in the application of terrestrial images for the generation of 3D point clouds in forests as a measurement tool for forest inventory (Piermattei et al., 2019, Mulverhill et al., 2019).

Many of the recent developments in terrestrial laser scanning can be of great interest in forestry (Liang et al., 2016). Fixed systems in tripods, have the limitation of reduced area coverage. Mobile systems, operated in a backpack and including an IMU for precise positioning in obstructed environments are an excellent solution. However, prices of such systems are high and probably not accessible for many users interested in forest inventory.

Photogrammetric systems are in general lighter, easier to transport and much more accessible. In particular, action cameras, commonly used in sports and outdoor activities, are very adequate for rough environmental conditions. The fact that many now have GPS units, makes them very interesting to do terrestrial photogrammetry in a forestry environment. This paper explores this fact, making use of an action camera and structure from motion (SfM) processing.

2. METHODOLOGY

The proposed method aims at obtaining dense clouds and 3D models, using the Structure from Motion (SfM) approach, with images acquired by an action camera in a forest environment.

The main purpose is to use the 3D model to measure dimensions of trees and other objects of interest within forestry inventory or general forest studies. The camera system to be used and the methodology are described in the following sections of the paper.

2.1 Description of the camera

A GoPro Hero 8 was used in this study. It can be mounted on top of a helmet and activated remotely by the operator with a smartphone. The operator controls the direction of the images simply by moving his head, looking at the objects of interest. The intention of the system is that in the field the operator just acquires image data, which will be later processed in the office. If much, only a reference tape measurement of some object may be done for quality assessment or eventual scale calibration (Mokroš et al., 2018).

The camera acquires video or discrete images. It includes a GPS receiver of navigation grade, with a positional accuracy of a few meters. Discrete images acquired in the JPEG format are tagged with WGS84 geographic coordinates and altitude above the ellipsoid, in the EXIF metadata. The positioning solution is an integration of GPS and attitude sensors, so even in a relatively dense forest environment, most of the images will have a position. Anyway there is a trend to a decrease in the accuracy inside dense tree coverage.

Action cameras are known for having a large radial distortion that increases very much the field of view. For a photogrammetric use it is preferable to use an alternative image mode, called “linear”, which corrects the essential component of the radial distortion. Figure 1 shows an image of a scene acquired in standard mode and in linear mode.

* Corresponding author



Figure 1. Example of GoPro images in standard mode (a) and linear mode (b).

A global correction model is applied to images of the linear mode. Each individual camera unit will probably have some residual distortion, so camera self-calibration will have to be taken into account in the processing. Analysing the images and their metadata it can be found that the images have, in the linear mode, the following characteristics:

Image size:	3000 lines by 4000 columns
Sensor pixel size:	1.73 μm
Focal length:	3 mm

The corresponding focal distance in pixel units is of 1733 pixels, but in the self-calibration it was always adjusted to values around 1905 pixels, which corresponds to an increase of 10% with respect to the starting value. Self-calibration was analysed and described below.

Another concern in terms of interior orientation is the effect of the rolling shutter. Although it is more noticed in video, image distortions can occur if the camera has a fast movement. Although an operator will in general have relatively small velocity when walking, this effect was also analysed. Common SfM softwares such as Pix4D Mapper and Agisoft Metashape include models for evaluating and correct the rolling shutter effect (Vautherin, 2016). Tests of camera calibration were done considering the rolling shutter correction turned on and turned off.

Images used in the tests described in this paper were acquired in the time lapse mode, with a time interval of 0.5 seconds. At the speed that an operator moves in the field, that image rate provides enough overlap for 3D data extraction. Images are recorded in JPEG format and tagged with the coordinates given by the GPS receiver. Alternatively, it would also be possible to acquire videos, at a much higher frame rate, and extract discrete frames to process in the very same way (Liu, et al., 2018). GPS positions are integrated in the MP4 video format and can be extracted to tag all the frames (Gonçalves and Pinhal, 2018). The advantage of acquiring discrete images is the higher resolution.

2.2 Description of the method

Images are acquired by the operator in an area covered by trees, with the camera pointing approximately in the horizontal direction, and following a path such that a set of trees will be covered from different directions. Tree trunks will show up in the images, in a way that for consecutive images there are common objects, not with very large changes of perspective. The operator should avoid passing very close to shrubs or other obstacles that produce very strong changes in the consecutive images, since that may cause problems for the image matching. From a set of images of trees, it should be possible to obtain a 3D model by SfM.

The main interest of the system is not the absolute geolocation accuracy of the 3D model but the accuracy of relative measurements such as tree height, or diameter. To start, we may think we have a set of images without any known geolocation. The image alignment in the SfM processing starts by identifying tie points in order to do a free bundle adjustment, in an arbitrary coordinate system (the model system). This results in the generation of an initial sparse point cloud, in the model system. This is similar to the concept of traditional relative orientation of a stereo pair of aerial photographs. A point with a position vector, \mathbf{u} , in the model coordinate system, is transformed into its georeferenced position vector, \mathbf{X} , by a 3D conformal transformation:

$$\mathbf{X} = s \cdot \mathbf{M} \cdot \mathbf{u} + \mathbf{T}_X \quad (1)$$

where \mathbf{M} is a rotation matrix, \mathbf{T}_X is a translation vector and s is a scale factor. If positions of camera projection centres are known, and are not collinear, they act as control points for absolute orientation of the model, using this equation. The point cloud will become georeferenced, but the accuracy of those positions will be similar to the accuracy of the camera GPS receiver, which is of a few meters.

The main concern in the system to be implemented is the accuracy of the scale factor, since it is critical for the ability of measuring object dimensions (Liang et al., 2014, Huang et al., 2018, Piermattei, et al., 2019). A standard procedure in terrestrial photogrammetry would be to measure one or more calibration distances. In the present case the idea is to depend, as much as possible, only on the GPS camera positions, avoiding calibration field measurements.

There is a risk that, if the accuracy of the projection centre GPS positions is low and they are in a small number, the scale factor may be very different from the reality, resulting in a low accuracy of the relative measurements. Increasing the number of images, the compensation effect of errors is expected to result in a more accurate scale factor. A first test was done with a small number of images, around an individual tree. Figure 2(a) shows a planimetric plot of the first case, in an area of 7 by 7 square meters, around a tree, where 31 images were used, with the positions given by the camera GPS receiver. Errors of a few meters in different directions originate a clear deformation of the surveyed path. After the image alignment, the positions become regularized, as can be seen on figure 2(b), revealing the actual shape of the path followed.

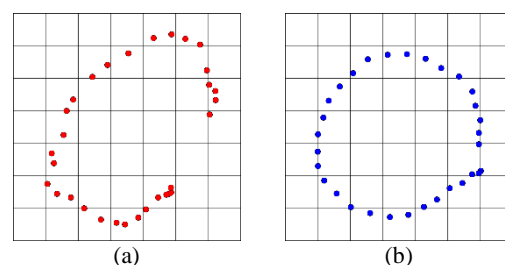


Figure 2. Horizontal camera positions: (a) as measured by the camera GPS, in red, and (b) after the SfM adjustment, in blue.

Figure 3 shows the generated dense point cloud and the measurement of a check distance of 1 meter. In the point cloud that distance showed up with an excess error of 30%. That is

explained by the large relative errors of the initial projection centre coordinates, which originated a poor quality solution of the absolute orientation.



Figure 3. Reconstructed model of the tree, with the measurement of a check distance.

In fact, looking at the orientation of the tree, we can see that it is tilted with respect to the Z axis of the reference system, while in fact it is vertical. That is explained by the height errors of the initial GPS coordinates. The plots of figure 2 show only planimetry, but the recorded camera heights had changes of 4 meters, and in fact the ground is flat and the camera was at a constant height. These errors in the camera projection centres introduced a tilt in the georeferenced model. Anyway, for the main purpose of the system the accuracy of the model location or orientation are not critical. The main point is to avoid a wrong scale, which would introduce a bias in the measurements to be done. The system will be tested with a much larger number of images, so that camera random errors in different directions tend to compensate.

3. TESTS OF THE SYSTEM

Two sets of tests were done at the Astronomical Observatory of the University of Porto, one for calibration purposes, in an open area, and a second set under tree coverage, in a situation closer to the intended real operating conditions of the method.

3.1 Tests in area without obstructions

The first set of tests were around the building of a telescope, which can be seen in figure 4. The area does not have obstacles for GPS positioning, the building has many details for image matching and some points are marked, which were measured with a total station. The main intention of these tests was to assess the camera calibration with different options and to assess the scale errors.



Figure 4. Building used for calibration tests.

The camera was taken around the building, at a slow speed, and collecting approximately 250 images. The first operation was to process these images in Agisoft Metashape, doing the image

alignment, with control points (marks on the building) and with camera self-calibration which resulted in what we called “LabC” parameters. The average RMS of reprojection errors was of 1.1 pixels. The parameters were, in the mathematical formulation used by Agisoft Metashape (Agisoft, 2022), the following:

Focal distance: 1904.5 pixels
Principal point (cx, cy): (10.7, 38.9) pixels
Radial (k1, k2, k3): (-0.00282, 0.00324, -0.00103)
Decentering (p1, p2): (0.00184, 0.00425)

Then a total of 5 sets (A1 to A5) like the first one, were collected and were processed without control points, with camera self-calibration and in two versions: without rolling shutter correction and then with rolling shutter correction. In all cases camera parameters were very similar, only with the difference that with rolling shutter correction, the principal point had larger variations of position. The average RMS of reprojection errors were always between 1.1 and 1.3 pixels. The main comparison of performance was made with the assessment of calibrated distances defined by the marked points. Figure 5 shows two of them. There is another marked distance of 15 m in the back of the building.

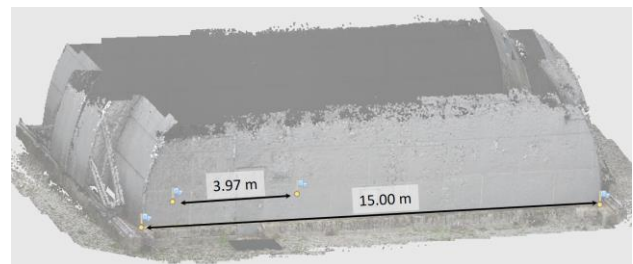


Figure 5. Building used for calibration tests.

The points were marked in several images and were then connected in order to compose scale bars. Their distances were evaluated according to the scale of the model. Since the scale only depends on the camera GPS coordinates, it is expected that the distances have some errors. Relative errors were calculated for the 15 m distances, for all the 5 sets, in the combinations: self-calibration (SelfC) with no rolling-shutter correction (RSC), SelfC with RSC and lab calibration (LabC) with no RSC. Table 1 contains the results. Additionally, the model scale was corrected with the 3.97 m distance acting as a control scale-bar (SB). It was applied to the 5 sets and the effect on the 15 m distances was assessed in the same way. The results are in the last column of table 1.

Test #	Calibration type			
	SelfC	SelfC+RSC	LabC	SelfC+SB
A1	1.4%	1.4%	1.3%	0.41%
A2	1.1%	0.8%	1.0%	0.24%
A3	3.4%	3.2%	3.4%	0.10%
A4	5.1%	6.0%	5.1%	0.08%
A5	5.1%	5.1%	5.1%	0.18%

Table 1. Relative errors in the length of verification distances of 15 meters.

Within each set, the results of the 3 combinations were very similar. It becomes evident that the effect of rolling shutter correction is not relevant. There is also no special advantage of using a lab calibration. In real operation of the system it will be

reasonable to always do the camera self-calibration, with no RSC.

The scale errors were small, around 1%, in some cases but rather large in others, around 5%, which may be acceptable in some situations, but might not be acceptable in others. To have more certainty of the accuracy, the use of a simple calibration bar may be useful. In all the tests the relative error became very small, well below 1%. Anyway, this test was done in very good conditions, with very well defined scale bars and with very rigorous distances. A real scenario in a tree environment will probably give lower accuracy. That was the reason for the second set of tests.

3.2 Tests under trees

Four tests were done under trees, in an area of the Observatory with evergreen oaks. These trees, although not composing a dense forest, are broad-leaved and cause a significant obstruction of GPS signal. They are the main trees of southern Portugal, and so this situation represents what can be found in many places where the system could be exploited.

Images were acquired in an area enclosed by a rectangle of 40 by 60 meters, along a path with approximately 300 meters, in a total between 450 to 500 images in each set. A levelling staff with an extension of 4 meters was placed on the ground in order to be used as a scale bar. Figure 6 shows one of the photos, where the staff can be seen. There were also some wooden buildings, which were used to assess the scale of the model.



Figure 6. Example of an image acquired in area covered by trees where the levelling staff is present

The set of 4 tests were identified as B1 to B4. Camera positions were recorded for all of the images, except in the case of B4, in which only around half of the images were tagged with coordinates. This is something that operators should be aware of in the field because the camera signals the correct functioning of the GPS. Figure 7 shows an aerial image of the area, with the camera horizontal GPS coordinates, which may look irregular because of the GPS accuracy.

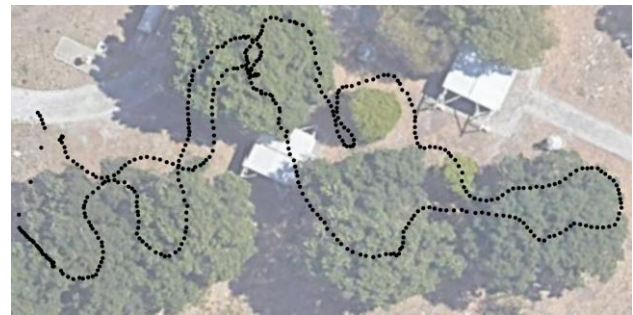


Figure 7. Positions of image acquisition in one of the sets.

The image alignment is done, with camera self-calibration, in order to obtain the complete orientation of the images. The points in the levelling staff and in the building walls are identified in the images where they are better seen.

Camera GPS positions and the adjusted ones after the SfM processing were plotted, in order to assess the changes. Figure 8 represents the two sets of camera positions: in red the original and in blue the new ones. Differences are of few meters, normally less than 3 m.

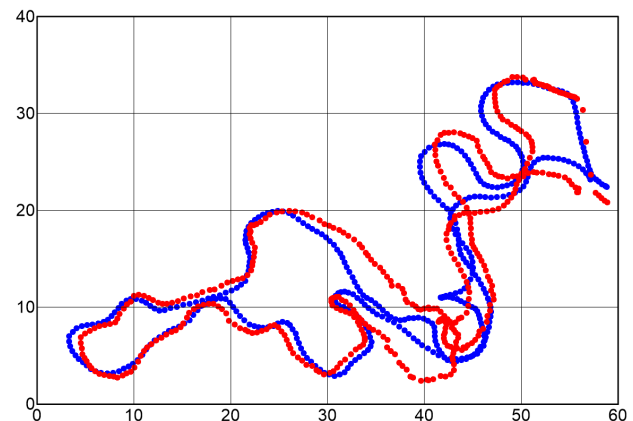


Figure 8. Original GPS camera positions (red) and corrected ones after SfM (blue). E and N axis in meters.

A similar analysis was done for the elevation, which was represented as a function of distance travelled. Figure 9 contains that graph, with distance and elevation in meters.

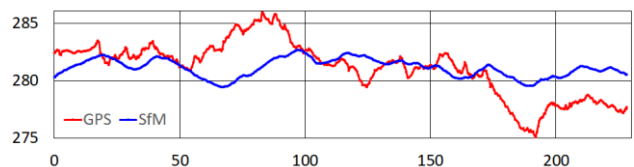


Figure 9. Original GPS camera elevation (red), and corrected ones after SfM (blue). Distance (x) and elevation (y) in meters.

Position changes occur in different directions, so it may be expected that there will be some compensation effect in the model scale.

Scale bars are created and the distance in the 3D model, d_{model} , is calculated. The relative error, R , is calculated with the known reference distance, $d_{reference}$, in the following way:

$$R_i = \frac{|d_{model} - d_{reference}|}{d_{reference}} \quad (2)$$

If several verification distances are considered in a set of images, the value to present is the average of the relative errors. In a second step the levelling staff was used as a calibration scale-bar, with 4 meters, and the average error in the other reference distances was calculated. Table 2 contains the relative errors for the 4 tests B1 to B4: in column “SelfC” are the data that had no model scale calibration (only interior orientation self-calibration); column SelfC+SB contains relative errors of the scaled models.

Test #	Calibration type	
	SelfC	SelfC + SB
B1	3.4%	2.0%
B2	2.4%	1.9%
B3	2.2%	0.6%
B4	6.6%	1.5%

Table 2. Relative errors of the 3D models obtained in the 4 image sets made under trees

There was an improvement in the model scale accuracy, when one scale bar is introduced, although not as significant as in the previous tests (A1 – A5). Figure 10 represents these data in bar chart. Set B4 had a very large error before calibration. The possible explanation: this was the case where only around half of the images had GPS positions. Images could be aligned and all ended up with coordinates, but the scale determined had a poor accuracy.

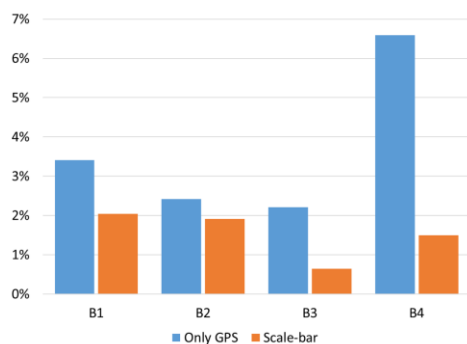


Figure 10. Representation of relative scale errors, with scale obtained by GPS only, and after the scale bar calibration.

Using a calibration scale bar, the model scale can be obtained with errors of between 1 and 2%. The system can be considered a reasonably good measuring tool for forests: 2% means an error of 2 cm in a distance of 1 meter, which is probably acceptable for many measurements in outdoor conditions.

Even depending only on GPS the error is relatively small. More tests will be done, in larger extensions and with more images to assess its performance and establish rules for the best use of the system.

4. CONCLUSIONS

The system developed for measuring dimensions of trees and other objects by terrestrial photogrammetry is very simple and cheap. It makes use of a GoPro camera, which is simple to operate and can be used in rough environments. The camera GPS unit provides positions even under trees, which helps obtaining a good georeferencing for most of the data.

The photogrammetric processing is done with Agisoft Metashape. It could be concluded that, on a regular basis, processing can be done with camera self-calibration and without rolling shutter correction.

The system provides point clouds and 3D models with an acceptable accuracy, which can be improved using scale bars. A scale bar of 1 or 2 meters can be easily transported and placed in the scene.

ACKNOWLEDGEMENTS

This work was developed within project 4Map4Health, financed by the Portuguese Foundation for Science and Technology, under program ERA NET CHIST-ERA.

The GNSS positioning was done with the ReNEP permanent stations of the Directorate General for Territorial Development (DGT).

REFERENCES

- Agisoft, L.L.C. (2022). Agisoft PhotoScan user manual: professional edition (version 1.8). St Petersburg, Russia: Agisoft LLC.
- Gonçalves, J. A., & Pinhal, A. (2018). Mobile Mapping System Based on Action Cameras. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, 42 (1), pp. 167-171.
- Liang, X., Jaakkola, A., Wang, Y., Hyypää, J., Honkavaara, E., Liu, J., & Kaartinen, H. (2014). The use of a hand-held camera for individual tree 3D mapping in forest sample plots. *Remote Sensing*, 6(7), 6587-6603.
- Liang, X., Wang, Y., Jaakkola, A., Kukko, A., Kaartinen, H., Hyypää, J., ... & Liu, J. (2015). Forest data collection using terrestrial image-based point clouds from a handheld camera compared to terrestrial and personal laser scanning. *IEEE transactions on geoscience and remote sensing*, 53(9), 5117-5132.
- Liang X., Kankare V., Hyypää J., Wang Y., Kukko A., Haggrén H., Yu X., Kaartinen H., Jaakkola A., Guan F., Holopainen M., Vastaranta M., (2016) Terrestrial laser scanning in forest inventories. *ISPRS Journal of Photogrammetry and Remote Sensing*, 115:63–77.
- Liu, J., Feng, Z., Yang, L., Mannan, A., Khan, T. U., Zhao, Z., & Cheng, Z. (2018). Extraction of sample plot parameters from 3D point cloud reconstruction based on combined RTK and CCD continuous photography. *Remote Sensing*, 10(8), 1299.

Mokroš, M., Liang, X., Surový, P., Valent, P., Čerňava, J., Chudý, F., ... & Merganič, J. (2018). Evaluation of close-range photogrammetry image collection methods for estimating tree diameters. *ISPRS International Journal of Geo-Information*, 7(3), 93.

Mulverhill, C., Coops, N. C., Tompalski, P., Bater, C. W., & Dick, A. R. (2019). The utility of terrestrial photogrammetry for assessment of tree volume and taper in boreal mixedwood forests. *Annals of Forest Science*, 76(3), 1-12.

Piermattei, L., Karel, W., Wang, D., Wieser, M., Mokroš, M., Surový, P., ... & Hollaus, M. (2019). Terrestrial structure from motion photogrammetry for deriving forest inventory data. *Remote Sensing*, 11(8), 950.

Vautherin, J., Rutishauser, S., Schneider-Zapp, K., Choi, H. F., Chovancova, V., Glass, A., & Strecha, C. (2016). Photogrammetric accuracy and modelling of rolling shutter cameras. *ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences*, 3(3).