Mobile Mapping System for Point Cloud Acquisition in a Forest Environment with an Action Camera

Andre Pinhal [1], Clara Lázaro [1,2], Jose A. Gonçalves [1,2]

[1] Astronomical Observatory, University of Porto, Monte da Virgem, 4430-146, Vila Nova de Gaia, Portugal
[2] DGAOT, University of Porto, Science Faculty, Rua Campo Alegre, 4169-007 Porto, Portugal
apinhal@fc.up.pt, clazaro@fc.up.pt, jagoncal@fc.up.pt

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

Abstract

This paper presents a methodology for generating georeferenced point clouds using a GoPro camera that incorporates a GNSS navigation receiver. The camera is mounted on a helmet worn by an operator, who simply needs to activate the camera and traverse the area of interest in overlapping strips, capturing 4K video at 60 fps. Frames are extracted from the video at an appropriate rate and processed using a structure-from-motion algorithm. This approach refines the camera's trajectory and produces a georeferenced dense point cloud. The georeferencing of the point cloud relies on the camera's GNSS-derived projection centres, which can be interpolated for each extracted frame. However, in forested environments, the reduced positional accuracy of the GNSS can introduce significant errors in the scale and orientation, limiting the accuracy of extracted dimensional parameters. To address these issues, the system incorporates a simple calibration strategy: a vertical pole of known length is placed in the surveyed area to provide a reference for scale and orientation correction. Once calibrated, the point cloud is processed to generate a canopy height model. Additionally, the point cloud can be segmented horizontally at 1.3 meters above ground level to extract individual tree rings, allowing measurements such as the diameter at breast height. The methodology is evaluated through various tests, and its accuracy is thoroughly analysed.

1. Introduction

Point clouds generated by mobile mapping systems (MMS) have a wide range of applications, with forest inventory being a particularly promising area of use (Mulverhill et al., 2019; Piermattei et al., 2019). The most commonly used systems for such applications are laser scanner-based (Liang et al., 2016). Several popular brands, including Riegl, Leica Geosystems, and Trimble, offer MMS solutions that integrate GNSS positioning, inertial navigation systems (INS), and data acquisition sensors such as cameras or laser scanners (Elhashash et al., 2022). These systems produce dense and highly accurate point clouds, even in forested environments. However, their costs, often reaching several hundred thousand euros, limit their accessibility to professional markets. This high cost makes such technology inaccessible to many users interested in forest surveys. Consequently, there is significant interest in the development of simpler, cost-effective systems based on photogrammetry and GNSS. These systems could provide, at a much lower cost, georeferenced point clouds capable of supporting forest measurements and tree identification (Mokroš et al., 2018; Pinhal and Gonçalves, 2022).

Recent advancements in photogrammetry, particularly those derived from computer vision methodologies, have greatly expanded its usability. The development of algorithms like Scale-Invariant Feature Transform (SIFT) (Lowe, 2004) significantly improved the efficiency of automatic conjugate point extraction between images, revolutionizing photogrammetric workflows. The combination of abundant automatic conjugate points from overlapping images and bundle adjustment techniques—adaptable to non-metric cameras via self-calibration—has democratized photogrammetry. This largely automated image orientation process, often referred to as "Structure from Motion" (SfM), enables processing of still images and video frames in diverse contexts, including aerial (e.g., UAV imagery) and terrestrial applications.

For video-based SfM, individual frames are extracted and treated as independent images. However, video frames are susceptible to distortions caused by the rolling shutter effect, which is not addressed by the traditional Brown camera model. Slow camera movement and high frame rates can mitigate this effect, and specialized rolling shutter correction models (e.g., Vautherin et al., 2016) are now integrated into leading photogrammetry software such as Pix4D Mapper and Agisoft Metashape.

SfM workflows typically begin with a free block adjustment that generates a sparse point cloud and estimates the camera projection centres in an arbitrary coordinate system. Georeferencing can be achieved using ground control points (GCPs) identified in the scene. However, for applications requiring rapid data acquisition, collecting GCP coordinates is impractical due to the time and effort involved, which undermines operational efficiency. Alternatively, cameras equipped with GNSS receivers can use their recorded positions to georeference the block via a seven-parameter transformation (translation, rotation, and scale). Action cameras like GoPro include built-in navigation-grade GNSS receivers, enabling position determination for both photos and video frames (GoPro, 2020). In open environments, these cameras typically achieve positional accuracy with a root mean square error (RMSE) of approximately 2 meters, as expected for navigation-grade GNSS devices (Petroskey et al., 2020).

In forest environments, GNSS performance is known to degrade significantly. While reduced positional precision is expected, the primary concern lies in the introduction of large errors in the camera's recorded positions, which can lead to critical scale

distortions in the 3D model. Such distortions severely impact the accuracy of tree dimension assessments and must be addressed.

When a point cloud is properly georeferenced, algorithms can be applied to automatically detect trees and generate models of their structure (Kükenbrink et al., 2022). Key metrics, such as the diameter at breast height (DBH) and tree volume, can then be extracted. DBH refers to the diameter of the tree trunk at a height of 1.3 meters above the ground. To derive this metric, ground classification algorithms are necessary to isolate points at ground level, enabling the creation of a height-filtered point cloud. By segmenting the point cloud at the 1.3-meter level, the resulting cross-sectional data can be analysed to identify closed rings, count them, and estimate the sectional area. However, these operations require a correctly oriented and scaled point cloud to ensure reliable measurements.

2. Description of the system implemented

GoPro cameras are widely used in outdoors to document recreational and sporting activities. Their robustness under challenging environmental conditions and ability to handle diverse lighting scenarios make them well-suited for the application described in this study. Additionally, their geolocation capability, which is of particular interest to users working with geospatial information, enhances their appeal. Starting with the GoPro Hero 5 model, these cameras have been equipped with GNSS receivers. For this work, the more recent Hero 8 Black model was utilized. However, a feature designed for sports users, the image stabilization, poses a challenge for photogrammetry as it dynamically shifts the camera's principal point. Therefore, this function must be disabled during operation.

In this study, video was chosen as the medium for image capture instead of standard discrete photos. Although the time-lapse mode allows for capturing images at short intervals (as little as 0.5 seconds), it may result in occasional coverage gaps. Video, in contrast, provides a much denser set of frames, allowing for improved overlap and ensuring continuous coverage.

The GoPro Hero 8 captures 4K resolution video, corresponding to frames of 3840 by 2160 pixels. The camera was set to a frame rate of 60 frames per second (fps), capturing one frame every 0.017 seconds. With its electronic shutter, each frame is acquired sequentially, line by line, over this time interval, rather than instantaneously. This sequential capture can introduce geometric distortions, known as the "rolling shutter" effect, caused by camera movement during acquisition. These distortions are not corrected by standard photogrammetric models, such as the Brown model. However, with an operator speed of 5 km/h, the displacement between consecutive frames is only about 2.4 cm, minimizing distortion. High frame rates further mitigate rolling shutter effects, though sudden rotational movements should still be avoided. While the software used for processing includes rolling shutter correction capabilities, the results obtained in all tests were satisfactory without enabling this option.

The camera was mounted on a helmet (Figure 1, left) worn by the operator, who acquired video while moving through a wooded area (Figure 1, right). The operator followed appropriate paths to create an overlapping scan of the area of interest for photogrammetric processing. While typically looking straight ahead, the operator may occasionally point the camera in other directions to enhance coverage. Sudden rotations must be avoided to minimize rolling shutter distortions.

Video recording was performed in "linear" mode, which corrects the strong radial distortions common in action camera lenses. The resulting frames have an equivalent focal length of approximately 1820 pixels, corresponding to a field of view of 93 degrees along the image's longest dimension (3840 pixels). This wide angle is equivalent to a 17 mm lens on a full-frame camera, classifying it as a very wide-angle lens.

Figure 1. GoPro camera mounted on top of a helmet (left), video recording by an operator in a wooded area (right).

To prepare for photogrammetric processing, multiple frames are extracted from the video based on the operator's movement speed, typically around 5 km/h. From tests conducted, extracting three images per second proved sufficient. These frames undergo an alignment process that generates a sparse point cloud and calculates the camera positions within an arbitrary reference system.

To georeference the entire block, the camera positions recorded during video capture are utilized. The GoPro video format is a MP4 file with embeded sensor data, including GNSS information, stored in the GoPro Metadata Format (GPMF). This metadata provides positional information (latitude, longitude, and height above the ellipsoid in the WGS84 system) and UTC time (in seconds of the day) for the start of each group of 60 frames. These data can be extracted using software tools provided by the manufacturer (GoPro, 2020). For photogrammetric processing, the geographic coordinates were converted to the Portuguese national map projection (EPSG:3763).

The extracted camera projection centres serve as control points for georeferencing. A 3D conformal transformation (Equation 1) is applied to adjust all points to a georeferenced coordinate system.

$$\boldsymbol{X} = S \cdot \boldsymbol{M}(\omega, \varphi, \kappa) \cdot \boldsymbol{u} + \boldsymbol{T_X}, \quad (1)$$

where $\boldsymbol{X}$ is the georeferenced position of a point, $\boldsymbol{u}$ is the position in the arbitrary system, $\boldsymbol{M}$ is a rotation matrix, dependent on 3 rotation angles, $(\omega, \varphi, \kappa)$, S is a scale factor and $\boldsymbol{T_X}$ is a translation vector, in the georeferenced reference system. The positional quality of the point cloud will depend on the quality of the camera positions, which may be degraded under forest cover. That was analyzed and improved as described in the following section.

3. Improvement of the point cloud

Several tests were conducted in a pine forest located in coastal dunes approximately 25 km south of Porto, Portugal. Three test sites, each covering an area of approximately 1,000 m², were surveyed. The sites were nearly flat, with maximum elevation variations of 3 meters, and characterized by dense tree canopies. Videos of about 4 minutes each were captured, with frames extracted at a frequency of 3 Hz. GNSS positional data were also extracted and interpolated to frame time. Initial image alignment was performed in Agisoft Metashape, without georeferencing,

enabling camera self-calibration. The stability of calibration parameters across tests was analysed, with results summarized in Table 1, which includes focal length (f) and principal point coordinates (c_x, c_y) in pixel units. These values were highly consistent, differing by at most 1 pixel. Although the frames were captured in "linear" mode, minor residual distortions of up to 6 pixels in the image corners were observed but remained stable across tests.

Test	f	c_x	c_x
1	1827.3	15.8	27.6
2	1827.0	15.7	28.6
3	1827.2	15.7	27.6

Table 1 – interior orientation parameters obtained in the self-calibration for the 3 tests analysed (pixel units).

There was initial concern about whether rolling shutter corrections would be necessary. However, the consistency of calibration parameters and the accurate epipolar lines predicted by Metashape, suggested that rolling shutter effects were negligible.

Next, camera positions were incorporated into the processing. GNSS data, extracted from the GoPro Metadata Format (GPMF), was given to Metashape. These positions were assigned low weights (10 meters a-priori standard deviation) in the adjustment process to allow for larger corrections of projection centres. After incorporating the GNSS data, a bundle adjustment was performed to refine the camera positions. Figure 2 compares the initial GNSS-derived trajectory with the adjusted trajectory. Initially, the path was irregular, but after alignment, it smoothed significantly.

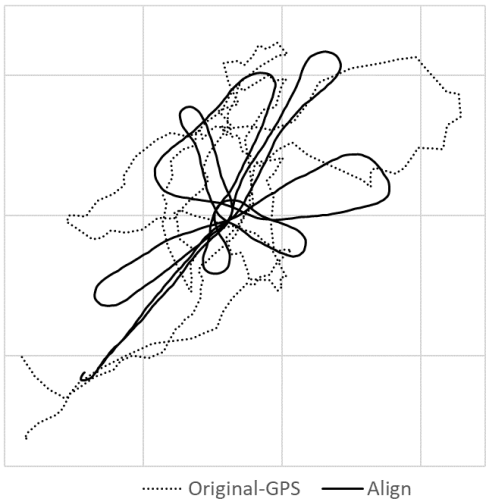

Figure 2. Camera trajectories: original, obtained by the GoPro navigation receiver (dashed line) and after the alignment (thick line). The grid has 10 metre spacing.

Figure 3 shows the altitude for both the original and adjusted trajectories. While the GNSS altitudes showed variations of around 50 meters, the corrected trajectory reduced these errors but still showed altitude variations much greater than the actual 1-meter elevation change in the test area. This error is caused by the point cloud tilt. Together with a probable scale distortion this point cloud is unsuitable for forestry measurements.

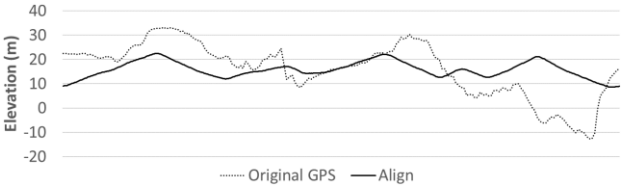

Figure 3. Camera altitude above sea level along consecutive frames: original, obtained by the GoPro receiver (dashed line) and after the alignment (thick line).

Ground control points (GCPs) could resolve these issues but would undermine the method's operational efficiency. Instead, a simpler strategy was implemented: a surveying pole was placed at the test site, stabilized with a bipod, and marked with precise top and base reference points. The distance between these marks, measured at 1.631 meters using a tape, served as a known scale. The pole was verticalized using a bubble level to allow for tilt corrections. Figure 4 illustrates the pole and its marked points in the aligned images.

Figure 4. Surveying pole, supported by a bipod (a), and marks in the top (b) and bottom (c) of the pole.

The pole's marks were manually identified in the images where they were visible with better definition. Their photogrammetrically estimated coordinates, (x_T, y_T, z_T) for the top and (x_T, y_T, z_T), for the base, were used to calculate the inclination angle, θ, relative to the vertical axis and the measured length, d, using equations (2):

$$\theta = \tan^{-1}\frac{\sqrt{(x_T - x_B)^2 + (y_T - y_B)^2}}{z_T - z_B} \qquad (2)$$

$$d = \sqrt{(x_T - x_B)^2 + (y_T - y_B)^2 + (z_T - z_B)^2}$$

Table 2 contains the results obtained for the 3 test sites: inclination angle, base-top distance and percentual error (relative to measured distance between the marks, 1.631 m).

Test site	θ (°)	d (m)	Relative error
1	30.2	2.595	59.2%
2	36.1	1.502	-7.8%
3	37.2	1.708	4.8%

Table 2. Vertical angle and distance of the line defined by the marks in the pole

Tests 2 and 3 showed moderate scale errors that exceeded acceptable limits for tree measurements. Test 1 exhibited a much larger error due to significant GNSS drift. Inclinations exceeding 30 degrees were also observed, indicating substantial deviations

from verticality. Both scale and tilt errors required correction to make the point clouds viable for forestry applications.

To address these issues, the pole's top and base points were treated as GCPs in a subsequent bundle adjustment. The base point was assigned its estimated coordinates, while the top point was given the same horizontal coordinates as the base, with an altitude 1.631 meters higher. These GCPs were assigned high weights (standard deviations of 1 mm), while the camera projection centres retained low weights (standard deviation 10 m). This adjustment does not improve georeferencing but corrects the scale and verticality, which are critical for horizontal tree sectioning and dimensional measurements. If approximate ground elevation values are available, they can also be incorporated for the base point to enhance altitude accuracy. However, azimuth correction can't be obtained, but it is not critical for the intended measurements.

Post-correction altitudes were plotted to assess improvements. Figure 5 shows the new camera elevations for test site 2. The height variation of only 0.9 meters reflects the site's nearly flat terrain.

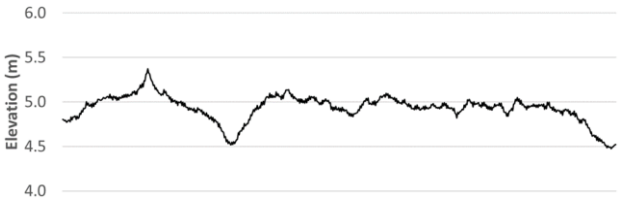

Figure 5. Camera elevations after correction (test site 2). The horizontal axis contains the successive camera positions.

To validate the newly generated point cloud, six wooden markers were affixed to trees within the survey area (Figure 6). These markers were precisely surveyed using a total station strategically placed within the forest, ensuring visibility of all markers. Outside the forest canopy, three points were surveyed using a differential GNSS receiver connected to a GNSS reference station. At these points, horizontal and vertical angle measurements were simultaneously recorded. Using the inverse resection method, the coordinates of the total station were calculated. Subsequently, the markers on the trees were measured using the total station's laser distance meter, with a positional accuracy estimated to be of 2 cm.

The six checkpoints were then identified in the Agisoft Metashape project on images where they were clearly visible. Their coordinates, as estimated photogrammetrically, were compared to those obtained from the total station survey.

Figure 7 represents the point cloud (in green), the check points and the pole, with radial lines from the pole (in red), as retrieved photogrammetrically. The points' true positions and radial lines are represented in blue. While the horizontal errors between the two datasets yielded a root mean square error (RMSE) of 14.7 meters, the relative geometry between the points was well-preserved. Horizontal distances between any pair of points differed by no more than 2 cm, confirming that the scaling and verticalization corrections were accurate. Despite georeferencing errors exceeding 10 meters and orientation discrepancies in azimuth, the sparse point cloud was deemed sufficient for extracting relative measurements critical for forestry applications.

Figure 6. Check points surveyed with a total station to assess the positional accuracy of data extracted.

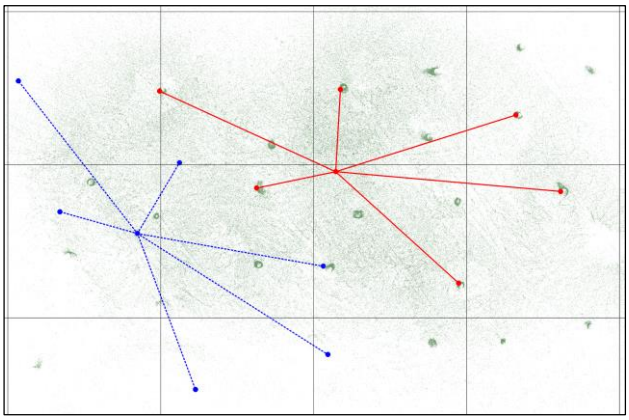

Figure 7. Planar representation of the point cloud, in green, pole and check points, in red, and their true (surveyed) positions, in blue. The grid spacing is 10 meters.

4. Extraction of tree parameters

The next step involved generating a "dense point cloud" using Agisoft Metashape. Medium density was chosen for this process, producing a 3D point for every 4×4 pixels in the original images (Agisoft, 2024). For the tests conducted with the GoPro camera, this configuration resulted in an average point spacing of approximately 1 cm in the dense cloud.

Dense point clouds enable the measurement of key forestry parameters, particularly the diameter at breast height (DBH), which is taken at 1.3 meters above ground level (Liang et al., 2016). Figure 8 (a) illustrates a dense point cloud derived from the image data. Due to the camera's ground-level perspective, typically oriented horizontally, only partial sections of the tree trunks are visible in the dense cloud, generally extending from the ground up to 5–10 meters. Despite this limitation, these sections are sufficient for detailed observations and accurate DBH measurements.

Figure 8(b) presents the point cloud with a color-coded altitude representation, highlighting variations in terrain elevation. On flat terrain, with a constant altitude, the point cloud can be directly sectioned. However, in undulating or uneven terrain, it is necessary to calculate the relative height of points above the ground surface to accurately isolate those corresponding to breast height.

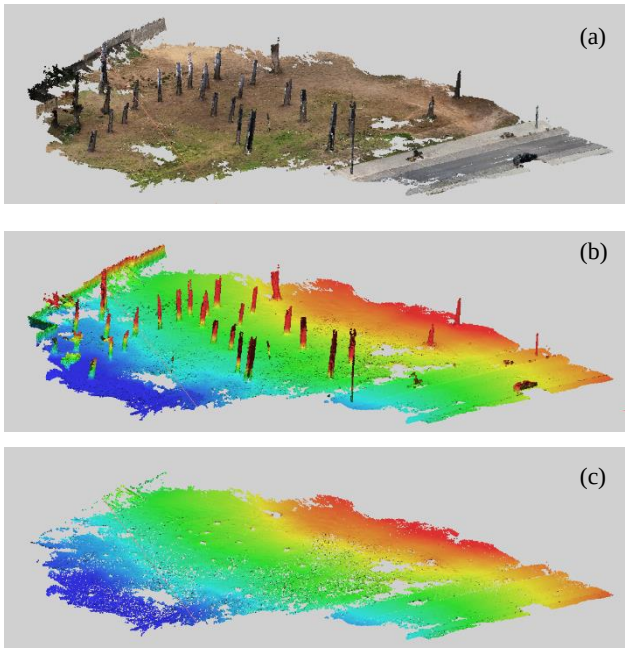

Figure 8. Dense point cloud (a), also represented with elevation coded by colours: (b) the full point cloud, including the trees, (c) point cloud after removing points classified as non-ground points

To address this, a ground point classification algorithm available in Agisoft Metashape was employed. This algorithm proved highly effective even in areas with some low vegetation. Figure 8(c) represents only the points classified as ground points. From this set of points a digital terrain model (DTM) was created. For every point in the dense cloud, the DTM's altitude value at that location was subtracted from the point's altitude, yielding the height of the point above the ground. This resulted in the creation of a canopy height point cloud (CHPC), which accurately represents the vertical structure of the surveyed area and facilitates detailed forestry measurements.

Finally, the CHPC is segmented under the condition 1.2 m<H<1.4 m, resulting in a subset of points concentrated on the tree trunks. These points are converted into a raster with 2 cm resolution. Subsequently, concave regions within the raster are filled, forming polygons. These polygons are then analysed to extract relevant data, including the number of detected trees. Figure 9 provides an example illustrating approximately 30 extracted trees.

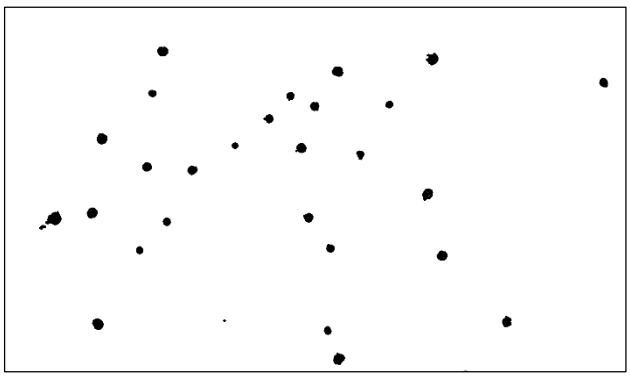

Figure 9. Image of approximately 30 trunks extracted in one of the test sites.

From the extracted polygons, areas are computed, and assuming a circular shape, the corresponding diameters are calculated. In the field, the perimeter at breast height can be measured using a tape, which is then converted to diameter under the same circular shape assumption. This process was performed for a set of large pine trees, and the results are presented in Table 3, showing the calculated DBH values in centimetres. Relative errors, expressed as percentages, were also calculated. The results appear to exhibit a bias, with point cloud measurements generally yielding smaller values than those obtained in the field. This discrepancy could suggest scale biasing or may be attributed to certain aspects of the adopted procedures. These issues will be examined in future system tests.

	DBH measurments (cm)		Relative
#	Point cloud	Field	error (%)
1	50.6	49.7	+1.9
2	50.6	52.5	-3.6
3	59.8	61.8	-3.1
4	66.2	68.8	-3.7
5	71.0	74.2	-4.3
6	73.2	76.4	-4.2
7	75.4	76.1	-0.8
8	78.0	80.2	-2.8
9	80.2	81.5	-1.6

Table 3 – Comparison of automatic DBH measurements in the point cloud and manual field measurements.

Special situations, such as forked trees below or above breast height, were not considered in this study. The method employed is relatively simplistic but remains practical and effective for tall trees, like the pine species analysed in this research.

5. Conclusions and future work

This study detailed a point cloud acquisition system for forest environments, using terrestrial photogrammetry based on video footage captured with a GoPro action camera mounted on a helmet worn by the operator. The system is user-friendly for field operators with no prior photogrammetry experience, as it requires adherence to straightforward guidelines, such as moving in consistent, regular strips and avoiding sudden camera movements.

The photogrammetric process generates georeferenced data utilizing the camera's GNSS receiver for positional information. However, due to significant positional errors under dense forest canopies, the resulting point clouds often exhibit scaling and inclination errors, rendering them unsuitable for precise tree measurements. A calibrated surveying pole with defined marks was introduced as a simple yet effective solution to these challenges. This approach corrected scaling and levelling errors, enabling accurate tree dimension measurements. Importantly, this ground control requirement involves minimal additional effort, as it only necessitates placing and levelling the pole within the survey area.

To extract the relevant data, a methodology was applied to remove the terrain altitude from the dense point cloud, creating a "canopy height model". Points corresponding to breast height (1.3 m, with a small tolerance) were extracted to automatically calculate tree diameters. These diameters were compared with field measurements in order to assess accuracy. On average, point cloud measurements underestimated diameters by approximately 2%. While the procedure requires further refinement, the error

margin is within acceptable limits for forest inventory applications.

The primary goal of this research was to establish and evaluate the methodology for point cloud generation using a GoPro camera. Although limited testing was conducted on automatic DBH assessment, the results are promising. Future experiments will aim to validate the methodology further, particularly under more complex conditions, such as varied terrain, increased low vegetation, and different tree species.

Acknowledgements

This work is part of the project 4Map4Health, under program ERA NET CHIST-ERA/0006/2019, funded by the Portuguese Foundation for Science and Technology (FCT).

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

References

Agisoft Development Team, 2024: Agisoft Metashape user manual: professional edition (version 2.1). St Petersburg, Russia: Agisoft LLC.

Elhashash, M., Albanwan, H., and Qin, R., 2022: A review of mobile mapping systems: From sensors to applications. Sensors 22.11 (2022): 4262. http://doi.org/10.3390/s22114262.

GoPro, 2020: GoPro Metadata Format – GPMF (v2.2.1). Processing software available at http://github.com/gopro/gpmf-parser.

Kükenbrink, D., Marty, M., Bösch, R., & Ginzler, C., 2022: Benchmarking laser scanning and terrestrial photogrammetry to extract forest inventory parameters in a complex temperate forest. International Journal of Applied Earth Observation and Geoinformation, 113, 102999. http://doi.org/10.1016/j.jag.2022.102999.

Liang X., Kankare V., Hyyppä 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. http://doi.org/10.1016/j.isprsjprs.2016.01.006.

Lowe, D. G., 2004: Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision, 60 (2004): 91-110. http://doi.org/10.1023/B:VISI.0000029664.99615.94.

Mokroš, M., Liang, X., Surový, P., Valent, P., Čerňava, J., Chudý, F., Tunák, D., Saloň, S., Merganič, J., 2018: Evaluation of close-range photogrammetry image collection methods for estimating tree diameters. ISPRS International Journal of Geo-Information, 7(3), 93. http://doi.org/10.3390/ijgi7030093.

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. http://doi.org/10.1007/s13595-019-0852-9.

Piermattei, L., Karel, W., Wang, D., Wieser, M., Mokroš, M., Surový, P., Koren, M., Tomaštík, J., Pfeifer, N., Hollaus, M., 2019: Terrestrial structure from motion photogrammetry for deriving forest inventory data. Remote Sensing, 11(8), 950. http://doi.org/10.3390/rs11080950.

Pinhal, A. and Gonçalves, J.A., 2022: Photogrammetric 3D Measurements in Forests with Action Cameras, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLIII-B2-2022, 449–454, http://doi.org/10.5194/isprs-archives-XLIII-B2-2022-449-2022.

Petroskey, Karla, Charles Funk, and Ivan A. Tibavinsky, 2020: Validation of Telemetry Data Acquisition Using GoPro Cameras. No. 2020-01-0875. SAE Technical Paper, 2020. http://doi.org/10.4271/2020-01-0875.

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 the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. III-3, 139-146. http://doi.org/10.5194/isprsannals-III-3-139-2016.