AUTOMATIC SINGLE PALM TREE DETECTION IN PLANTATIONS USING UAV-BASED PHOTOGRAMMETRIC POINT CLOUDS

T. Kattenborn\textsuperscript{a} , M. Sperlich\textsuperscript{b} , K. Bataua\textsuperscript{a} , B. Koch\textsuperscript{a}

\textsuperscript{a} FeLis, Chair of Remote Sensing and Landscape Information Systems, University Freiburg, Tennenbacherstr. 4, 79106 Freiburg, Germany - ferninfo@felis.uni-freiburg.de

\textsuperscript{b} SOPAC, Applied Geoscience and Technology Division of Secretariat of the Pacific Community - SPC, 241 Mead Road, Nacula Suva, Fiji Islands - director@sopac.org

Commission VI, WG VI/4

KEY WORDS: Single Tree Detection, UAV, Palm plantation, Structure From Motion, Point Clouds, Segmentation, Terrain Models

ABSTRACT:

For reasons of documentation, management and certification there is a high interest in efficient inventories of palm plantations on the single plant level. Recent developments in unmanned aerial vehicle (UAV) technology facilitate spatial and temporal flexible acquisition of high resolution 3D data. Common single tree detection approaches are based on Very High Resolution (VHR) satellite or Airborne Laser Scanning (ALS) data. However, VHR data is often limited to clouds and does not allow for height measurements. VHR and in particular ALS data are characterized by high relatively high acquisition costs. Sperlich et al. (2013) already demonstrated the high potential of UAV-based photogrammetric point clouds for single tree detection using pouring algorithms. This approach was adjusted and improved for an application on palm plantation. The 9.4ha test site on Tarawa, Kiribati, comprised densely scattered growing palms, as well as abundant undergrowth and trees. Using a standard consumer grade camera mounted on an octocopter two flight campaigns at 70m and 100m altitude were performed to evaluate the effect Ground Sampling Distance (GSD) and image overlap. To avoid commission errors and improve the terrain interpolation the point clouds were classified based on the geometric characteristics of the classes, i.e. (1) palm, (2) other vegetation (3) and ground. The mapping accuracy amounts for 86.1% for the entire study area and 98.2% for dense growing palm stands. We conclude that this flexible and automatic approach has high capabilities for operational use.

1. INTRODUCTION

1.1 Palm plantation in the global context

Rising global demands for natural resources induce an increase in intensive land use forms such as palm plantations (Mekhilef et al. 2011; Koh & Wilcove 2007). The latter, e.g. oil or coconut plantations supply a wide variety of products, ranging from food, biofuels, construction timber, firewood, cosmetics or textile fabric (Koh & Wilcove 2007; Ohler, 1999). For reasons of planning and documentation there is a high interest in efficient inventories on the individual tree level (Shafri et al. 2011). Furthermore, due to rising concerns about sustainability and global change industrial players and stakeholders aim for a traceable certification of the production (Basiron 2007). However, with respect to the large extents of those plantations terrestrial inventory methods are time consuming and therefore expensive.

1.2 Remote Sensing for Plantation Management

Thus, for plantation inventories remote sensing data and analysis techniques are valuable tools. A few studies have used optical very high resolution (VHR) satellite imagery to manually count palms or automatically segment palms crowns based on their spectral and textural characteristics (Shafri et al. 2011; Korom et al. 2014; Kamiran & Sarker 2014). These studies achieved high classification accuracies (75-95%). However, VHR data is characterized by relatively high acquisition costs, and data acquisition on a regular basis is often limited by cloud cover. Furthermore, conventional VHR imagery does not allow for height measurements of palm individuals or surrounding vegetation as shrubs or trees. LIDAR offers precise and dense 3D data and is therefore widely and commercially used in forest inventories (Hyyppa et al. 2008). Shafri et al. (2012) analyzed the potential for airborne LIDAR and optical data for plantation planning and inventory. Despite the highly accurate 3D and optical data this technology is relatively expensive for inventories with a high temporal resolution, especially with regard to plantations in remote regions or islands.

Recent developments in Unmanned Aerial Vehicle (UAV) technology facilitate spatial and temporal flexible acquisition of high resolution optical data. Costs and required expert knowledge are constantly decreasing due to fast advancements in the development of hard- and software (Colomina & Molina 2014). By applying photogrammetric processing techniques, such as structure from motion algorithms, UAVs also allow for a hypetemporal and hyperspatial data acquisition of 3D point clouds (Luceeer et al. 2012).

With regard to automated inventories on the individual plant level many point cloud based detection methods have been studied and developed for Aerial Laser Scanning (ALS) data (Vauhkonen et al. 2012, Kartinen et al. 2012), Sperlich et al. (2013) already demonstrated the high potential of UAV-based photogrammetric point clouds for an automatic single tree delineation in deciduous and coniferous forest stands. The applied pouring algorithm, originally developed for ALS data, showed particularly high detection accuracies in the presence of distinct tree tops. We thereby hypothesize that this 3-dimensional segmentation technique will provide high detection accuracies for palm individuals and the measurement of their crown widths and heights. In addition we aim to improve the individual detection by a preprocessing of the data, to additionally prevent other vegetation being classified as palms.
2. METHODOLOGY

2.1 Study Site and UAV Data Acquisition

The test area is situated on Tarawa, an atoll in the Republic of Kiribati, in the central Pacific Ocean. The test site, with a size of approximately 470*200m, was chosen, since it comprises dense and sparse coconut palm cover as well as abundant varieties of undergrowth and other trees (see fig.1). Using this composition we aim to evaluate the detection accuracy of the reference palms as well as the robustness of palm detection in presence of other vegetation.

As UAV-platform an octocopter, MK Okto2 (Highsystems), was chosen. In contrast to fixed-wing UAV-platforms multicopters allow for a high image overlaps and carriage of high quality cameras. As sensor a consumer camera (Panasonic Lumix G3), with a 20mm lens was mounted on the UAV using a gimbal. In order to estimate the effects of ground sampling distance (GSD) and image overlap, two flights were carried out at different altitudes of 70m and 100m. The autonomous UAV-flights were pre-configured using 4 parallel stripes. The camera direction was to nadir, recording single frames (ISO 400, shutter speed 1/1200s) at a frequency of 1.4Hz. The two data sets were acquired consecutively to ensure nearly equal illumination and wind conditions. A summary of the photogrammetric characteristics of the two data sets is given in Table 2.

<table>
<thead>
<tr>
<th>Alt.</th>
<th>GSD (cm)</th>
<th>forward overl.</th>
<th>side overl.</th>
<th>images</th>
</tr>
</thead>
<tbody>
<tr>
<td>70 m</td>
<td>1.3</td>
<td>89%</td>
<td>32%</td>
<td>519</td>
</tr>
<tr>
<td>100 m</td>
<td>1.9</td>
<td>93%</td>
<td>58%</td>
<td>501</td>
</tr>
</tbody>
</table>

Table 1: Average ground sampling distance (GSD), forward and side overlap as well as the total number of images for both data sets.

2.2 Point Cloud Generation and Reference data Collection

The cleaned set of images was processed with a structure from motion tool chain using VisualSFM (V0.5.24, Furukawa & Ponce 2010). In order to increase reconstruction completeness and quality we calculated the internal camera parameters using the calibration software 3DF Lapyx (3DFlow). Accordingly a fixed camera calibration was used during the VisualSFM image matching step. The 70m and 100m datasets were co-registered using 8 checker boards as artificial ground control points, which were evenly spread within the study site during the flight campaigns. To ensure the comparability of the two data sets the resulting point clouds were clipped to a joint spatial extend.

In respect to the size of the test area, and in order to include as many individuals as possible the reference data was collected from the point clouds itself. Reference palm positions were identified visually aided by the respective image frames. This procedure also allowed to simultaneously check the local quality of the reconstruction. In the rare case of incomplete reconstructions the presence of apparent palm cast shadows and the single image frames allowed for an accurate identification of the palm position. Furthermore the height of each individual was measured using the top elevation of the palm crown within a normalized Digital Surface Model (nDSM, see chapter 3).

After Co-registration and reference data collection we decided to reduce the point cloud density for further processing and analysis. A voxel-grid filter with a minumum distance of 0.1m was applied to both datasets (70m and 100m).

2.2.1 Thematic Point Cloud Classification

To exclude non-palm vegetation during the delineation process the point cloud was classified into three classes: (1) Palm (2) other vegetation including shrubs and trees and (3) ground including bare soil and grass. Besides reducing omission errors during the palm tree detection this approach is also meant to enable the mapping of undergrowth and other vegetation to evaluate the accessibility and assist cultivation management. Furthermore a segmentation of ground points is likely improve the interpolation of a Digital Terrain Model (DTM), required for the application of the pouring algorithm.

The RGB information of the reconstructed point clouds did not allow for an accurate separation of the desired classes on spectral basis. Hence, we focused on the geometric characteristics of
each class to separate the single points. The classification was carried out using the multi-scale dimensionality criterion classification developed by Brodú & Lague (2012), which was originally designed for Terrestrial Laser Scanning (TLS) data. In brief, the algorithm identifies the local dimensionality characteristics of the point cloud using neighborhood balls at different scales, i.e. search radii. On this basis a classifier is constructed, which describes the best combination of scales to separate the user-defined categories on the basis of their geometric difference. For the training of the classifiers we manually segmented 30 palm individuals of the 3D-reconstruction and approximately 200m² of both other vegetation and bare ground. For both data sets training and application of the classifiers was carried out separately. The classifiers were trained using 43 different scale factors respectively diameters between 0.1m - 6m. Initially the upper limit of 6 m was set as the processing time exponentially increases with larger scale factors. Accordingly we performed a step-wise exclusion of the largest scales in order to evaluate the potential accuracy gain of the latter.

2.3 Palm Individual Detection and Accuracy Assessment

The palm detection was performed using the software TreesVis (Weinacker et al. 2004) through the implemented pouring algorithm. First, on the basis of the processed point clouds DTM and DSM were calculated in order to generate a nDSM, i.e. vegetation heights. Then, using the pouring algorithm palm tops and the according crown boundaries were estimated in the nDSM. Similar to a watershed algorithm the pouring algorithm calculates local maxima of the nDSM aided by a gaussian filter. The maxima are the starting point for a downhill directed sequence, searching for the minima surrounding the maxima. However, as these minima do not always correspond to the edge of the crown a ray algorithm adjusts the crown border for a more realistic delineation. Thereby crown border points are corrected, if distances and angles of virtual rays between the tree tops end the surrounding minima exceed predefined thresholds.

Different parameters, i.e. Gaussian Filter size, DSM and DTM interpolation methods, as well as height constraints were adjusted to palms on a heuristic basis. For accuracy assessment and validation a 2m search radius was applied to the reference tree positions. A modeled palm was considered as detected, if it was within the specified radius of a reference tree. As a primary quality criterion of the palm detection we calculated the overall Mapping Accuracy (MA), based on detected reference palms, non detected palms (omission error) and falsely detected palms (comission error):

\[
MA = \frac{\text{detected}}{\text{detected} + \text{omission} + \text{comission}} \tag{1}
\]

For further evaluation of the modeled palms, we measured the euclidean distances between reference and modeled palm position (x,y) as well as the height deviations (z) for each individual.

3. RESULTS

3.1 3D Reconstruction Quality

Bare ground and other vegetation show highest point densities and completeness. Since the images were recorded in nadir-direction most of the densely growing palms lack stem points due to occlusion by the overlapping crown. The reconstructed palm crowns of the 70m data set tend to be slightly less noisy. In contrast the 100m data set shows a higher degree of reconstruction completeness in terms of crown representation and total number of reconstructed palms. Out of 615 palms within the study area the number of reconstructed palms sums up to 479 (77.9%) for the 70m data set and 597 (97.1%) for the 100m data set. Thereby a palm was defined as not reconstructed, if only a few points of the crown or only the parts of the stem were reconstructed.

3.2 Thematic Point Cloud Classification

The multi-scale dimensionality criterion classification produced unexpectedly good results and allowed for a clear separation of the point clouds into the three classes. Only negligeable portions of the reconstructed palms were classified as other vegetation see (fig. 4). Very few points of other trees or tall growing shrubs were classified as palms. We assumed that these few points would either be ignored during DSM interpolation or would not meet the geometric characteristics to be delineated as a palm during the application of the pouring algorithm. Falsely classified ground points were aggregated in relatively small patches. A subsequent clustering of the ground point cloud allowed for a removal of clusters with spatial extend lower than a manual selected threshold (0.5 m).

The stepwise exclusion of the largest scales (see fig.2) revealed that the classification accuracy of palms and other vegetation increases with larger scales. The proportion of falsely classified palm points decreases almost linear with larger scale factors. Falsely classified shrub points show a significantly stronger decrease of falsely classified points with larger scales. Both the shrub and palm classification accuracy show a tendency to converge at scales larger than 6m. Based on the described relationship we did not expect major classification accuracy gains adding larger scales (>6m), which would significantly increase computation durations. The classification of the ground class showed the overall highest accuracy. For all scales the classification accuracy exceeds 95%.

Figure 2: Proportion of falsely classified palm and shrub points (%) for different scale ranges. For each of the classifications a step-wise exclusion of the largest scales was performed. For visual interpretation a loess fit was calculated for both the palm and shrub classification results. The shown graph is based on the 100m dataset.

3.3 Tree detection

For the 70m and 100m dataset best performance was archived using equal parameters for the pouring algorithm. The gaussian filter size was set to 20 pixels (nDSM pixel size = 0.25 m). Since the smallest reference tree was 3.9 m high the minimum height contraint for the pouring algorithm was set to 3.5 m. As described...
in chapter 3 neither within the 70m nor the 100m point cloud all reference trees were defined as reconstructed. Thus, we analyzed the detection results based on all reference trees in order to include both the photogrammetric reconstruction and pouring algorithm inaccuracies.

Corresponding to the total number of reconstructed palms the overall mapping accuracy (MA) amounts for 68.6% for the 70m and 86.1% for the 100m data set respectively (see table 2). Indicated by the high omission errors (27.6%) the low MA of the 70m data set is caused by its lower number of total reconstructed palms.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>data set</th>
<th>70m</th>
<th>100m</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. Palms</td>
<td></td>
<td>615</td>
<td>615</td>
</tr>
<tr>
<td>No. Palms reconstructed</td>
<td></td>
<td>479</td>
<td>597</td>
</tr>
<tr>
<td>No. Palms Modelled</td>
<td></td>
<td>499</td>
<td>589</td>
</tr>
<tr>
<td>No. Palms Detected</td>
<td></td>
<td>445</td>
<td>557</td>
</tr>
<tr>
<td>Omission Error (%)</td>
<td></td>
<td>27.6</td>
<td>9.4</td>
</tr>
<tr>
<td>Comission Error (%)</td>
<td></td>
<td>10.8</td>
<td>5.4</td>
</tr>
<tr>
<td>Mapping Accuracy (%)</td>
<td></td>
<td>68.6</td>
<td>86.1</td>
</tr>
</tbody>
</table>

Table 2: Palm Tree Detection results for the two datasets (70m and 100m) based on all reference palms.

Each model included the estimated crown margins, the position (x,y) and the height (z) of the palm (see fig.4). The averaged position and height deviations of the reference and modeled palms are listed in table 3. Both deviations of the estimated parameters are almost equal for both data sets. In general the heights are underestimated due to the filtering of single vertically exposed palm leaves (measured palm height) by the pouring algorithm. Although most of the palms in the study area show a tilted growth habit the average position deviation is clearly below the specified tolerance radius of 2 m.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>data set</th>
<th>70m</th>
<th>100m</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean Δ height (m)</td>
<td></td>
<td>-0.44</td>
<td>-0.52</td>
</tr>
<tr>
<td>mean Δ position (m)</td>
<td></td>
<td>0.61</td>
<td>0.59</td>
</tr>
</tbody>
</table>

Table 3: Averaged height and position deviations between reference and modeled palms for both data sets (70m and 100m).

Reference trees and detection results were mapped in order to analyze the detection accuracy spatially (see fig. 3). Fewest omission and comission errors are present in the dense palm areas, where only 6 individuals could not be detected. The corresponding MA in this area is 98.2%. More frequent omissions occur at rather isolated palm individuals. Most comission errors occur in the presence of abundant shrub and tree cover, where small fragments of the point cloud remained after the classification process.

4. DISCUSSION

Within the present study we developed a work flow for an automatic palm plantation inventory, based on UAV-based photogrammetric point clouds. The processing of these point clouds includes a geometric classification in order to ensure the detection of the target class, i.e., palm individuals. The following sections will discuss the results and findings with regard to the components of the presented methodology.

4.1 Data acquisition and 3D reconstruction

The results indicate that an increased forward and side-overlap (at higher altitudes) improves the overall 3D reconstruction of the palm plantation in terms of reconstruction completeness. More frequent overlaps increase the probability of 3D reconstruction of detected features, such as palm leaves. This was even pronounced within the present study, as palm leaves are easily displaced by wind and gusts, which compounds the feature matching. Accordingly the more frequent overlapping of the 100m data set raised the chance of acquiring multiple images of a palm leave in the same position. We assume that there are optimization capabilities within the structure from motion processing. Accordingly different matching and reconstruction parameters have to be adjusted to palm vegetation. Further campaigns have to be conducted to identify the best ratio of GSD and side-overlap. In general flight campaigns should be performed in preferably calm wind conditions.

Using an UAV-based LIDAR could most likely overcome effects of wind and gusts. Yet, these systems are costlier and have a shorter operating duration due to heavier weight of the payload.

4.2 Thematic Classification

The applied classification procedure, originally developed for TLS data, was successfully applied to the photogrammetric point cloud. First, a removal of non palm vegetation reduces commission errors and therefore increases overall mapping accuracy of the palm detection. Second, the classification allows to identify and quantify undergrowth for managing purposes. By that a height model would allow for an easy interpretation of vegetation presence and
intensity. The latter could ease cultivation activities such as regular clearances or for maintenance of the road network. Third, the classification of ground pixels ensures that only the latter were included during the DTM interpolation. We assume that especially dense and extended shrubs would have impaired the interpolation results. As shown in figure 3, depending on the geometric characteristics of the class, the classification accuracy strongly depends on the value range of the scales. The overall highest accuracy of the ground class is a result of the simple and distinct geometry, where points adjoin in a relatively flat and almost 2D-plane. Therefore larger search radii of the neighbourhood ball do strongly enhance the classification results. In contrast the geometric characteristics of palms and shrubs are explained and separated at a broader range of scales. Hence, not only the geometric structure of adjacent points is important, but also the overall form of leaves and branches. We assume that the beginning of the classification accuracy convergence (<4.5m) is linked with the diameter of the palm crowns. At this scale also visual interpretation is most likely to differentiate between the geometric pattern of shrubs and palm crowns. Furthermore, larger neighbourhood balls and respectively larger search radii, would include large proportions of both shrub and palm points without explaining the geometric characteristics of a single class.

Due to the overall unexpectedly robust classification results we also expect a high potential of the multi-scale criterion classification of UAV-based photogrammetric point clouds in other disciplines, such as archaeology, forestry or land-cover mapping.

4.3 Single Tree Detection

We demonstrated that the pouring algorithm is a sound method for detecting palms and measuring geometric variables, such as height or crown with. The marked-off crown top of palms enables an accurate detection in densely vegetated areas. In general omission errors are a consequence of incomplete palm crown reconstructions during the structure from motion processing. Accordingly rather outlying palms, readily exposed to wind and gusts,
featured a relatively high proportion of omission errors. Fritz et al. (2013) presented a methodology using UAV derived photogrammetric point clouds in which a RANSAC-based cylinder fit is used to detect single tree stems. This will most likely feature a valuable extension of the present approach, since the outlying trees are also more likely to have their stems reproduced in the point cloud. As our study site comprises other vegetation, such as dense undergrowth and abundant trees the results are not readily comparable to other methods. However, with an overall MA of 86.1% within the 100m data set the proposed methodology features a high detection performance. With regard to the automatic procedure the MA of 98.1% in dense growing palm stands is unexpectedly high. The overall position of the modeled palm trees show a very high correspondence with the reference trees. To a certain extent height deviations are indisputable, since the absolute height of an individual often relies on a single palm leaf. The general underestimation of the palm tree height is caused by the filtering of the pouring algorithm, which filters vertically protruding points, e.g. single top leaves. Yet, we assume that a palm height measurement is only possible to decimeter accuracy, since the defined top palm leaf height can possibly fluctuate according to the water status and therefore vigor. With an adjustment of the pouring algorithm the measured heights will enable a valuable predictor for classifying age classes or estimate biomass as well as carbon stocks using available allometric functions (Asari et al. 2013).

5. CONCLUSION
Within this study we could confirm the a high potential of UAV-based photogrammetric point clouds for single tree detection on plantations. Overall, the proposed approach can be regarded as a highly competitive remote sensing solution for palm plantation inventories. The cost effective, flexible and mobile UAV-technology and the highly automatized processing chain can be considered as available for operational use.

ACKNOWLEDGEMENTS
The authors would like to thank Prof. M. Petterson and Dr. W. Forstreuter (SOPAC - Applied Geoscience and Technology Division of Secretariat of the Pacific Community) for technical support, Dr. G. Kattenborn (Gecopter.eu) for providing the UAV and equipment, the Ministry of Environment, Lands and Agriculture Development, Kiribati as well as the Ministry of Fisheries, Minerals and Resources, Kiribati for on site assistance.

REFERENCES