

Original papers

Determination of stem position and height of reconstructed maize plants using a time-of-flight camera



Manuel Vázquez-Arellano^{a,*}, Dimitris S. Parafos, David Reiser^a, Miguel Garrido-Izard^b, Hans W. Griepentrog^a

^a Institute of Agricultural Engineering, University of Hohenheim, Garbenstraße 9, 70599 Stuttgart, Germany

^b Laboratorio de Propiedades Físicas (LPF)-TAGRALIA, Technical University of Madrid, Madrid 28040, Spain

ARTICLE INFO

ABSTRACT

Keywords:
Kinect v2
Crop characterization
Agricultural robotics
Precision farming
Plant phenotyping

Three dimensional (3-D) reconstruction of maize plant morphology by proximal sensing in agriculture brings high definition data that can be used for a number of applications related with precision agriculture and agricultural robotics. However, 3-D reconstruction without methodologies for extracting useful information is a senseless strategy. In this research, a methodology for stem position estimation is presented relying on the merging of four point clouds, using the Iterative Closes Point algorithm, that were generated from different 3-D perspective views. The proposed methodology is based on bivariate point density histograms for detecting the regional maxima and a radius filter based on the closest Euclidean distance. Then, single plant segmentation was performed by projecting a spatial cylindrical boundary around the estimated stem positions on a merged plant and soil point cloud. After performing a local Random Sample Consensus, the segmented plant point cloud was clustered using the Density-based spatial clustering of applications with noise algorithm. Additionally, a height profile was generated by rasterizing the plant and soil point clouds, separately, with different cell widths. The rasterized soil point cloud was meshed, and the rasterized plant points to soil mesh distance was calculated. The resulting plant stem positions were estimated with an average mean error and standard deviation of 24 mm and 14 mm, respectively. Equivalently, the average mean error and standard deviation of the individual plant height estimation was 30 mm and 35 mm, respectively. Finally, the overall plant height profile mean error average was 8.7 mm. Thus it is possible to determine the stem position and plant height of reconstructed maize plants using a low-cost time-of-flight camera.

1. Introduction

One of the most appealing aspects of reconstructing the geometry of an agricultural environment is to obtain information about the crop status without the troublesome manual measurement. Doing so with an efficient investment in resources, such as fuel and working time (Steckel, 2017), would trigger the interest of farmers in this technology. The information provided by the scanned and digitized data would be very useful for decision-making throughout the cropping cycle; considering that it involves precision agriculture practices (Gebbers and Adamchuk, 2010). Tasks such as sensing and mapping: soil, crop, weed and yield are suitable for 3-D imaging systems. Other applications such as agricultural robotics and plant phenotyping for breeding purposes are among the most appealing (Blackmore et al., 2006). However, since research using 3-D imaging systems in agriculture was previously

limited, particularly with the once expensive time-of-flight (ToF) cameras, there is still the need of new methodologies for extracting useful information out of the 3-D data for agricultural applications. Information such as stem diameter, plant height, leaf angle, leaf area index (LAI), number of leaves, biomass, etc. are of particular interest. If the cost of obtaining such information becomes economically accessible, new applications and solutions will come as a result (Vázquez-Arellano et al., 2016a). An off-the-shelf ToF camera such as the Kinect v2 (Microsoft, Redmond, WA, USA) offers a good cost/performance ratio solution for the development of ground-based 3-D imaging system for proximal sensing (Vázquez-Arellano et al., 2016b). In 3-D imaging a point cloud is a set of data points in space, where each point $P(x, y, z)$ is a function of the spatial position (x, y, z) in a Cartesian coordinate system (Rusu and Cousins, 2011).

Until now, among the most commonly measured plant parameters

* Corresponding author.

E-mail addresses: mvazquez@uni-hohenheim.de (M. Vázquez-Arellano), d.parafos@uni-hohenheim.de (D.S. Parafos), dreiser@uni-hohenheim.de (D. Reiser), miguel.garrido.izard@upm.es (M. Garrido-Izard), hw.griepentrog@uni-hohenheim.de (H.W. Griepentrog).

using this ToF camera in agricultural research are plant height and biomass estimation. Recently, Hämerle and Höfle (2016) developed a mobile system for maize plant height measurement using a Kinect v2 where they used a terrestrial laser scanner (3-D LIDAR) to digitize maize plants (as the reference) from different perspectives, using real time kinematic-global navigation satellite system (RTK-GNSS) for georeferentiation, and artificial markers to facilitate the point cloud registration and alignment. With the Kinect v2 they obtained depth information, and through raster crop height model with a rough cell resolution ($1\text{ m} \times 1\text{ m}$), they approximated the maize height with an R^2 determination coefficient of 0.89 with one of the approaches. However, they acknowledged that the accuracy was slightly below the results of other studies due to the rough terrain and the complex maize architecture, among others. Andújar et al. (2016) used the same ToF camera for weed volume estimation with R^2 determination coefficient for weed biomass of 0.7 but 0.58 for maize biomass. The lower value for maize could be explained by its complex architecture since one single perspective cannot describe it entirely, compared to the low-lying weed. Ribeiro et al. (2017) reconstructed vineyards using a small electric car and the Kinect v2 ToF camera with an RTK-GNSS for geo-referentiation. They relied on a variant of the ICP algorithm for the point cloud registration and stitching to reduce the problem of drifting. Then, they developed a four-step methodology for segmenting the canopy points from the entire point cloud. Finally, they used alpha shapes to envelop the canopy point clouds in order to create a volume map of the vineyard rows.

Individual maize plant phenotyping was also investigated by Lu et al. (2017) where they developed a robotic arm 3-D imaging acquisition system based on a SR-4000 ToF camera (MESA Imaging, Rueschlikon, Switzerland). They obtained measurements of different phenotypic traits such as stem height, leaf length and angle, and number of leaves of individual plants on pots. A similar research was done by Chaivivatrakul et al. (2014) using the same SR-4000 ToF camera but with the plant pot placed on a turntable driven by a stepper motor. They also achieved stem and leaves segmentation and phenotypic data extraction such as stem diameter; and leaf length, area and angle. They mentioned that the most challenging parameter, and with the highest error (21.89%), was the leaf area due to partial occlusions and rolling of some leaves. Aside from that, they also used the non-uniform rational basis spline algorithm for surface reconstruction for a 3-D holographic visualization. Nakarmi and Tang (2012) used a ToF camera to measure the maize inter-plant spacing by mosaicking depth images using encoder readings and a feature matching algorithm. They achieved an overall root mean square error (RMSE) of 0.017 m and a misidentification ration of 2.2%, concluding that the camera position of their research (side-view) achieved superior accuracies compared with previous researches (top-view) for inter-plant spacing sensing.

The aim of this research was to estimate the stem position of maize plant point clouds, calculate the height of the individual plants and generate a plant height profile of the rows using a low-cost ToF camera. In order to validate the stem position estimations with the real world, the seedling positions were used as ground truth data (on-site measurement using a total station after plant emergence), and for the plant height validation, manual measurements were used. The main contribution of this research is the estimation of single plant position and height to evaluate the potential and limits of the ToF camera used in the research.

2. Materials and methods

2.1. Hardware and sensors

The 3-D data used in this research was obtained using a robotic platform, developed at the University of Hohenheim, controlled by a joystick that navigated between maize plants in a greenhouse. The ToF camera was mounted at the front of the vehicle at a height of 0.94 m

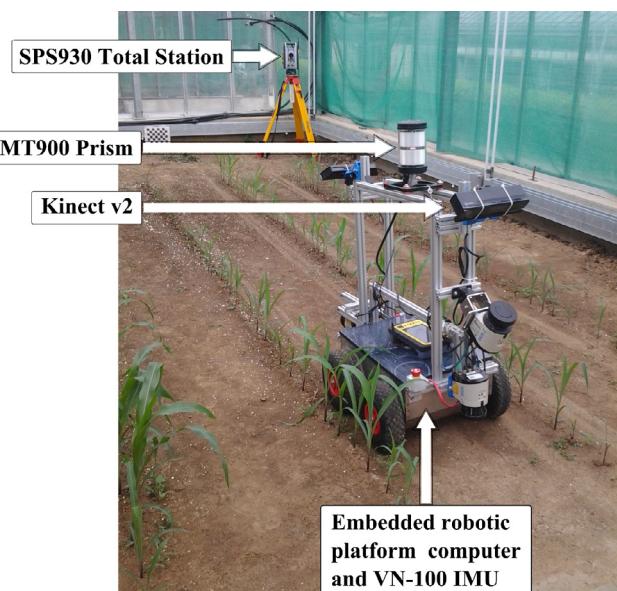


Fig. 1. 3-D imaging acquisition system with the TALOS robotic platform. The different components for 3-D image, orientation and position data acquisition are marked with arrows and annotations.

with a downwards angle of 45 degrees. The SPS930 robotic total station (Trimble Navigation Limited, Sunnyvale, USA) tracked the position of the robot by aiming at the Trimble MT900 Machine Target Prism. An Inertial Measurement Unit (IMU) (VectorNav, Dallas, USA) VN-100 was embedded inside the robotic platform and used to measure its orientation while driving. The 3-D imaging acquisition system is depicted in Fig. 1 and the technical characteristics of the robotic platform are described in detail by Reiser et al. (2015).

2.2. Experimental setup

The experiment was done in a greenhouse ($3.75\text{ m} \times 5.6\text{ m}$) at the University of Hohenheim. The maize was planted in 5 rows with different standard deviations from the theoretical spacing: the inter-row spacing was 750 mm and the intra-row spacing was 130 mm. This deviation during seeding was done in order to emulate different seeding scenarios. From row 1 to 5 the standard deviations were 19, 17, 6, 48 and 47 mm, respectively. Every row had 41 plants in a length of 5.2 m, and the plant growth stage was between V1 and V4 (Ritchie et al., 1992). However, most of the plants (94%) were between V1 and V3. The ground truth was measured with a robotic total station tracking the target prism, mounted on a tripod, and pointing directly over each seedling with the help of a plummet. The robot platform was driven, using a joystick, in every path in the go and return direction. At every headland, the robot was turning 180 degrees, therefore, the 3-D perspective view was different in the go and return direction of every row. A viewpoint was established (camera plot in Fig. 2), to avoid confusion between the left and right side of the crop row. Since there were 3-D reconstructions while going and returning, the left and right side would be different (if a viewpoint was not established) depending on the driving direction. Additionally, the experimental setup is represented and depicted in Fig. 2.

2.3. Data processing

The raw data for this research was based on the maize plant registration and stitching from a previous research (Vázquez-Arellano et al., 2018). These point clouds were processed mainly using the Computer Vision System Toolbox™ of MATLAB R2016b (MathWorks, Natick, MA, USA). Also, some functionalities of CloudCompare (EDF R&D, 2011)

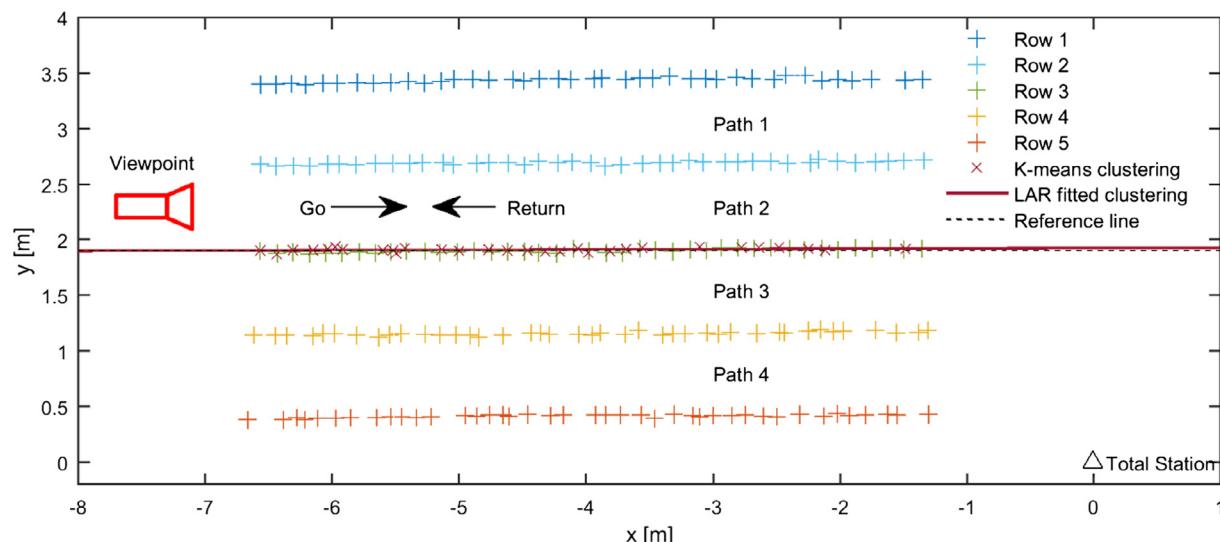


Fig. 2. Depiction of the greenhouse maize seedlings (+) and the *k*-means clustering (x) with the least absolute residuals (LAR) representative fitting and reference lines. The viewpoint is depicted as a camera plot used to avoid confusion between left and right side of the crop row.

were used for point cloud processing. In this research, only row 2, 3 and 4 were analysed; row 1 and 5 were discarded since they were scanned from just one side, in Fig. 2 it can be seen that they are both near the edges of the greenhouse where the robotic platform did not fit.

2.3.1. Crop row alignment

In the research by Vázquez-Arellano et al. (2018), an approximation of the plant stems was done using the *k*-means clustering at a thin layer of 5 mm at the bottom of the plants. The main intention was to evaluate the accuracy of the reconstruction by relating the clusters with the seedlings positions (ground truth), but also to know the alignment of the reconstructed crop rows by fitting a first order polynomial curve (Eq. (1)) considering the least absolute residuals (LAR) to minimize the effect of extreme values on the fit; were p1 and p2 were 0.01382 and 1.908, respectively.

$$f(x) = p1*x + p2 \quad (1)$$

This line provided a representative angle of the crop rows relative to the x axis of the total station coordinate system. During the data acquisition, the x axis of the total station coordinate system was intentionally aligned as accurate as possible with the crop rows, in order to facilitate the data processing. However, in open field conditions the crop rows are rarely aligned with the coordinate system of the georeferencing system. Therefore, the LAR fitting line was used to transform the reconstructed point cloud by rotating it until it was aligned with the x axis. In Fig. 2 it can be seen that the angular difference between the LAR fitting of row 3 and the reference line (parallel to the x axis) is minimal (0.0046 degrees). The row 3 clustering was used as a representation of the whole maize rows since it was in the middle row.

2.3.2. Plant stem position estimation methodology

The methodology for stem position estimation proposed in this research relied on reconstructed crop rows using the data and methodology proposed by Vázquez-Arellano et al. (2018). The generated point clouds were used to obtain data point density histograms using a bivariate approach by pairing the x and y values of every point in the point cloud (see Fig. 3b). The registration and alignment of the point clouds was done pairwise, relying on the ICP algorithm, with two crop row point clouds at a time (generated from different perspective views) until all of them (four in total) were aligned. They were later merged together into a single point cloud and filtered resulting in a crop row point cloud like the one depicted in Fig. 3a. The bivariate point density histogram depicted in Fig. 3b shows the number of points (frequency of

occurrence) that fall into each bin of size 1 cm². Thus, the bigger the plant the larger the frequency, but more important, the closer to the stem position the larger the frequency. This concept can be visualized in another perspective if the crop row point cloud is seen from above in the x-y axis (see Fig. 3c) where the bins of warmer colour represent a higher frequency (of occurrence) and thus a higher point density. This is closely related with the stem position since the stem points are concentrated in the center of the plant and the leaf points are scattered around it.

Previous research, such as the one by Lu et al. (2015), used univariate point density histograms for detecting the stem of a single plant from different 3-D perspective views. They considered the stem of the plant as the mean value of the histogram plus 3 times the standard deviation. This approach of using univariate point density histograms works well for single plant stem segmentation but it is not suitable for multiple plant analysis. The main reason is that with the univariate approach in a maize row, the information about the position of the individual stems in one axis (x axis in Fig. 4) is provided as a form of a local peak. But in the perpendicular axis (y axis in Fig. 4) that information about the position of the individual stems would not be provided, because the y axis value would be the same for all plants since the histogram has a single peak value for the whole maize row. In order to know the position of the individual stems, a unique x and y value must be provided.

In single plant morphological analysis, a precise plant stem estimation is an early step in the pipeline to extract other plant parameters, however, in high throughput morphological analysis it is often avoided, rather estimating overall parameters such as biomass or height profiles, due to the difficulty of its detection. This difficulty relies on the fact that plant stem estimation is not an aim by itself, rather an objective to reach the main aim which is single plant segmentation. Single plant segmentation requires complex 3-D imaging algorithms (Reiser et al., 2018) or even machine learning. In this research, a precise plant stem estimation was considered as a prerequisite for further plant morphological analysis. For instance, in the case of maize, it is difficult to automatically obtain the stem diameter, segment the leaves or calculate the height if the stem position was not previously determined. The methodology for stem position estimation proposed in this research is shown in the flowchart of Fig. 5.

The process started by importing a point cloud pair (a) of two rows, later, filtering each point cloud with a radius outlier removal (ROR) filter and statistical outlier (SOR) filter- their input parameters are shown in Fig. 5. Then, computing the Random Sample Consensus

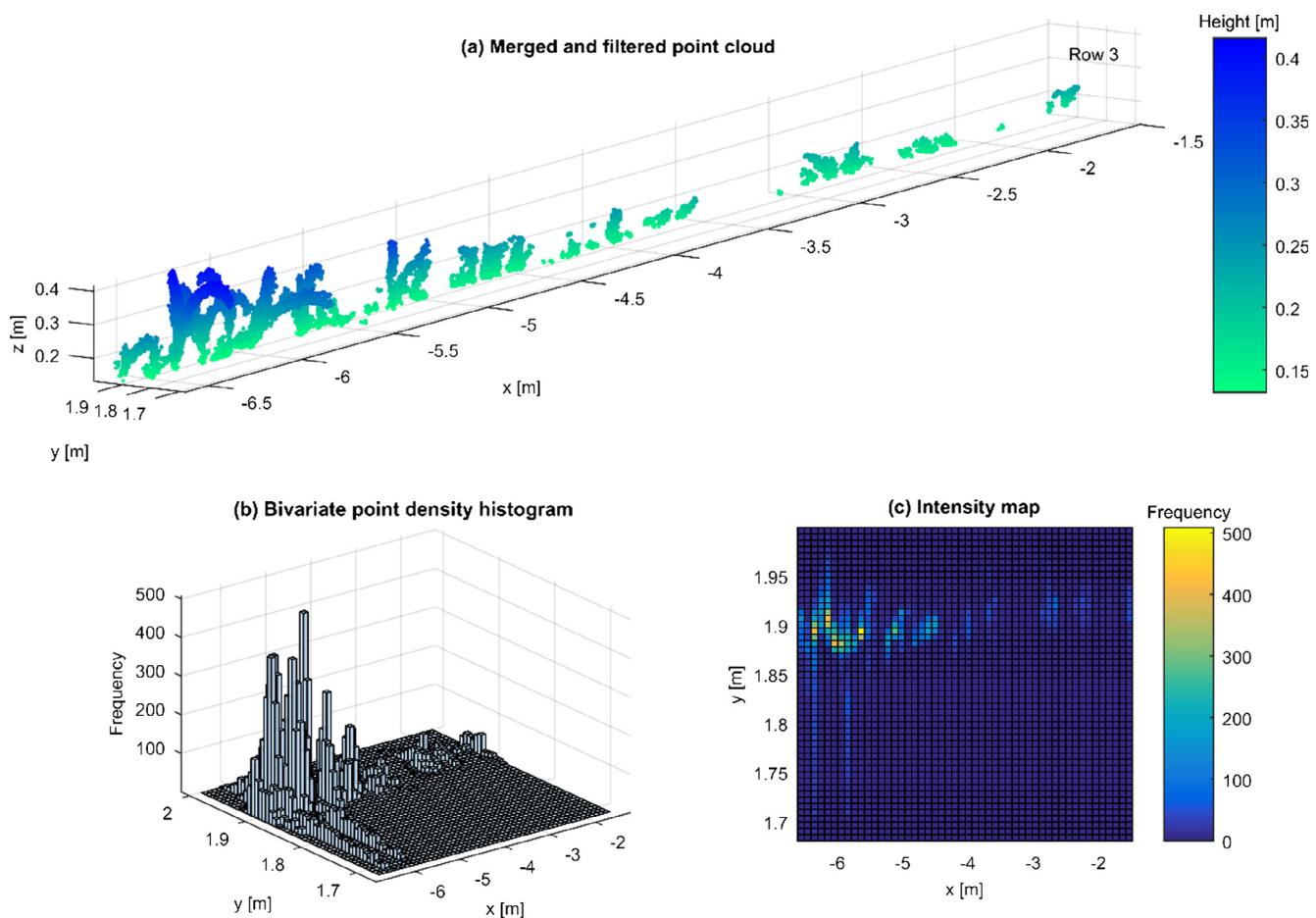


Fig. 3. Result of the registration and alignment of four maize row point clouds after (a) merging and filtering. (b) The bivariate point density histogram of the resulted maize row point cloud. (c) Another representation, as an intensity image, where the warmer squares indicate a local maxima.

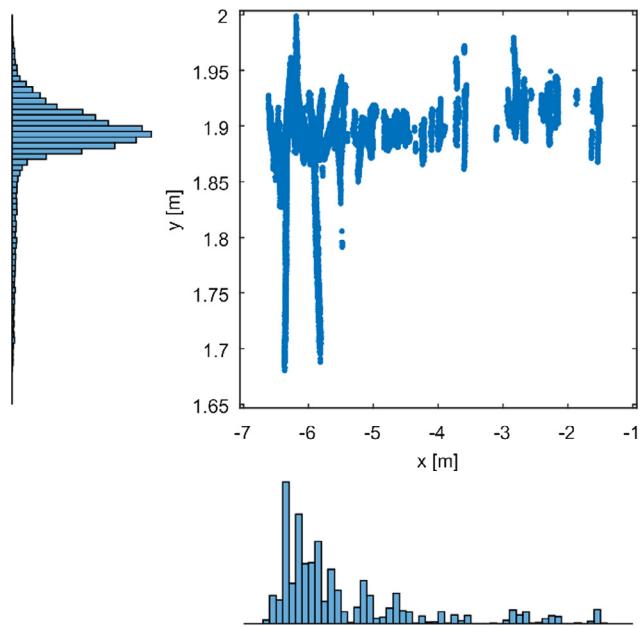


Fig. 4. Univariate point density histogram in the x and y axis of the point cloud depicted in Fig. 3a.

(RANSAC) algorithm (Fischler and Bolles, 1981), for each point cloud pair, with the maximum distance from an inlier to the plane set to 0.5 times the theoretical distance between the plants (65 mm). Next,

The disadvantage of merging several point clouds was that the errors (e.g. due to the point cloud assembly, ToF camera accuracy, moving leaves) accumulate. It was noticed that after merging 4 point clouds, some of the plants appeared thicker or slightly blurred as they were in reality. Therefore, an assessment of the accuracy of the point cloud alignment was necessary to provide an idea of how well they overlapped. For this, the Cloud-to-Cloud (C2C) method (Girardeau-Montaut et al., 2005) was used to measure the quality of the overlapping by measuring the average distance of one point cloud compared with another one set as a reference, it is also used for detecting changes, therefore this method can be useful for plant growth monitoring. For every point in the compared point cloud, the closest point was defined in the reference point cloud, then the absolute distance was computed and the mean and standard deviation of all the point distances was calculated. CloudCompare software was used to compute the distances and to calculate the mean and the standard deviation. The greater the distance error, the more different the point clouds were. However, it was not possible to distinguish the source of the difference, which could be due to errors in the point cloud generation, moving plants, 3-D perspective view and other parameters.

After the aligned point clouds were merged, a subsequent subsampling was applied using a voxel grid (3 mm × 3 mm × 3 mm) filter, and then, a SOR (20 pts. 1 nsigma) filter for noise removal, where $nsigma$ means n times the standard deviation. At this point, most of the plants point clouds were a continuum without flying (veil) points

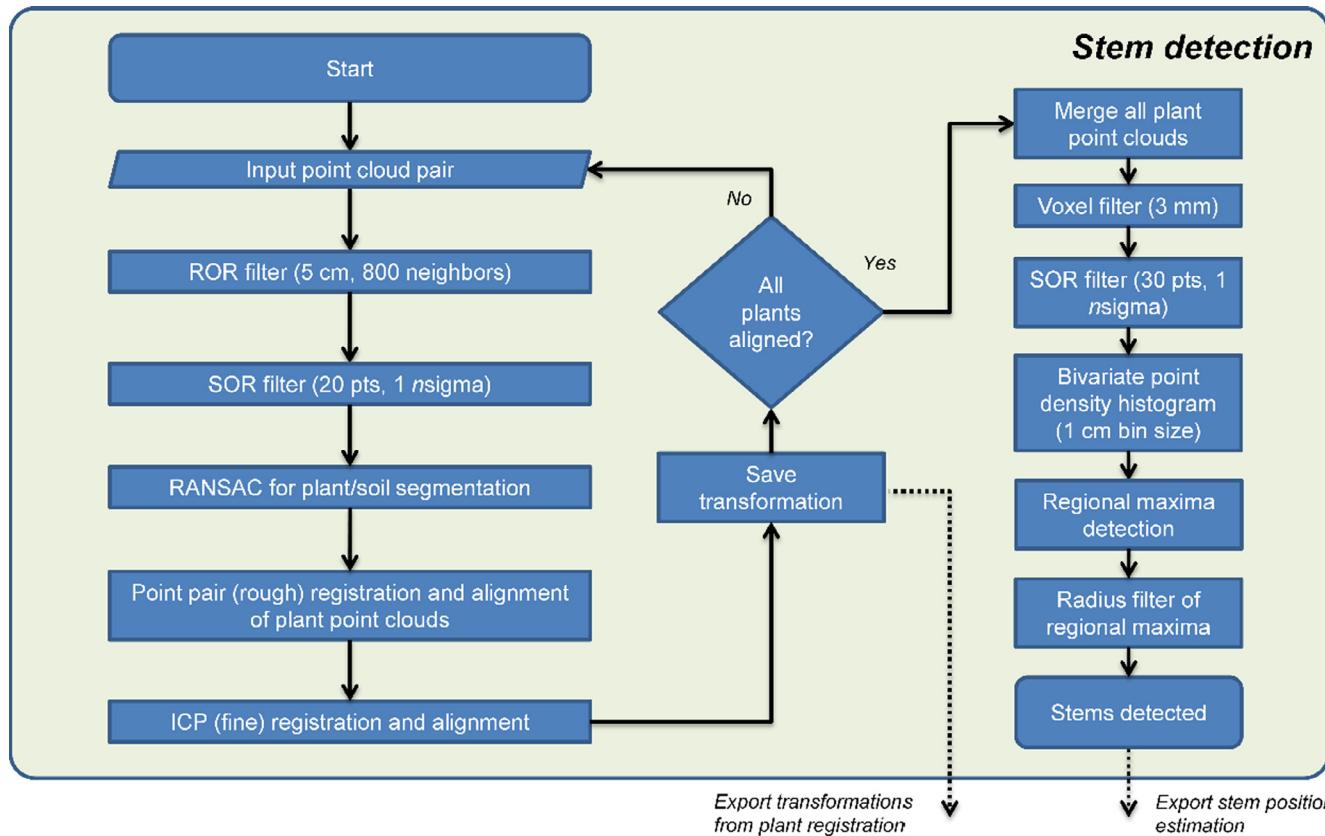


Fig. 5. Stem detection flowchart showing the pipeline process starting with the input point cloud pair, through the filtering and merging, until the stem detection based on bivariate point density histograms.

(Steder and Konolige, 2011) orbiting them. However, the main disadvantage of applying noise reduction filters was that the ToF camera was able, in some cases, to obtain 3-D data of small plants, but since the density and cohesion of those points was low, they were removed by the filter. After filtering, the remaining plant points were used to create a point density histogram with a bin size of 1 cm^2 . The resulting bivariate point density histogram was plotted using the shape-preserving Piecewise Cubic Hermite Interpolating Polynomial (PCHIP) method (Kahaner et al., 1989) for interpolation.

After plotting the point density histogram, the next step was to detect the regional maxima. The regional maxima were connected components of bins with a constant value, t , whose external boundary bins all had a value less than t . To compute the connectivity, the algorithm uses 8-connected neighborhoods, in a 3-by-3 matrix, for every bin.

To remove the regional maxima outliers, a filtering based on the radius filter was implemented. The regional maxima with the largest point count, in other words: the tallest peak, was considered as the starting reference point, since it was the one with the most probabilities of being a stem. The next stem position was then estimated in one direction of the x axis, by adding 130 mm, and the maximum regional maxima searching distance was established with Eq. (2).

$$\text{max_dist} = \text{spacing_factor} * \text{plant_spacing} \quad (2)$$

In this research the maximum distance was $13 \text{ mm} * 3 = 39 \text{ mm}$, therefore, the spacing factor was 3 and the plant spacing 13 mm. The regional maxima points lying within the maximum distance (max_dist) radius were sorted by the shortest Euclidian distance to the expected stem position. Then, the closest points were filtered out if their count was bigger than a threshold (set to 20 points). After all the regional maxima points were analysed in one direction of the x axis, the reference point was once again returned to the regional maxima with the

largest point count, in order to perform the same procedure but in the opposite direction. The disadvantage of merging several point clouds was that the roughness of the plant surface and edges increases, due the differences between them, thus increasing the count variability within neighbouring bins resulting in more detected (noisy) peaks. The main part of the functionality of the filter is better described with [Algorithm 1](#), where $p_i.X$ and $p_i.Y$ are related with the bivariate approach previously mentioned.

Algorithm 1. Radius Filter

```

Require: pointCloudPlants, nextPlantPosition, maxDist
for all points p in pointCloudPlants do
    dx = p_i.X - nextPlantPosition.X
    dy = p_i.Y - nextPlantPosition.Y
    point_dist = euclideanDistance (dx, dy)
    if point_dist < maxDist then
        pointCloudRadiusInliers (p_i) = PointCloudPlants (p_i)
    end if
end for
Require: scan, sensorOrigin, minAngle, maxAngle
for all points p in scan do
    angle_to_last_point = getAngleBetweenPointsComparedToSensorOrigin (p i, pi - 1)
    angle_to_next_point = getAngleBetweenPointsComparedToSensorOrigin (p i, pi + 1)
    if angle_to_last_point > minAngle or
        angle_to_last_point < maxAngle then

```

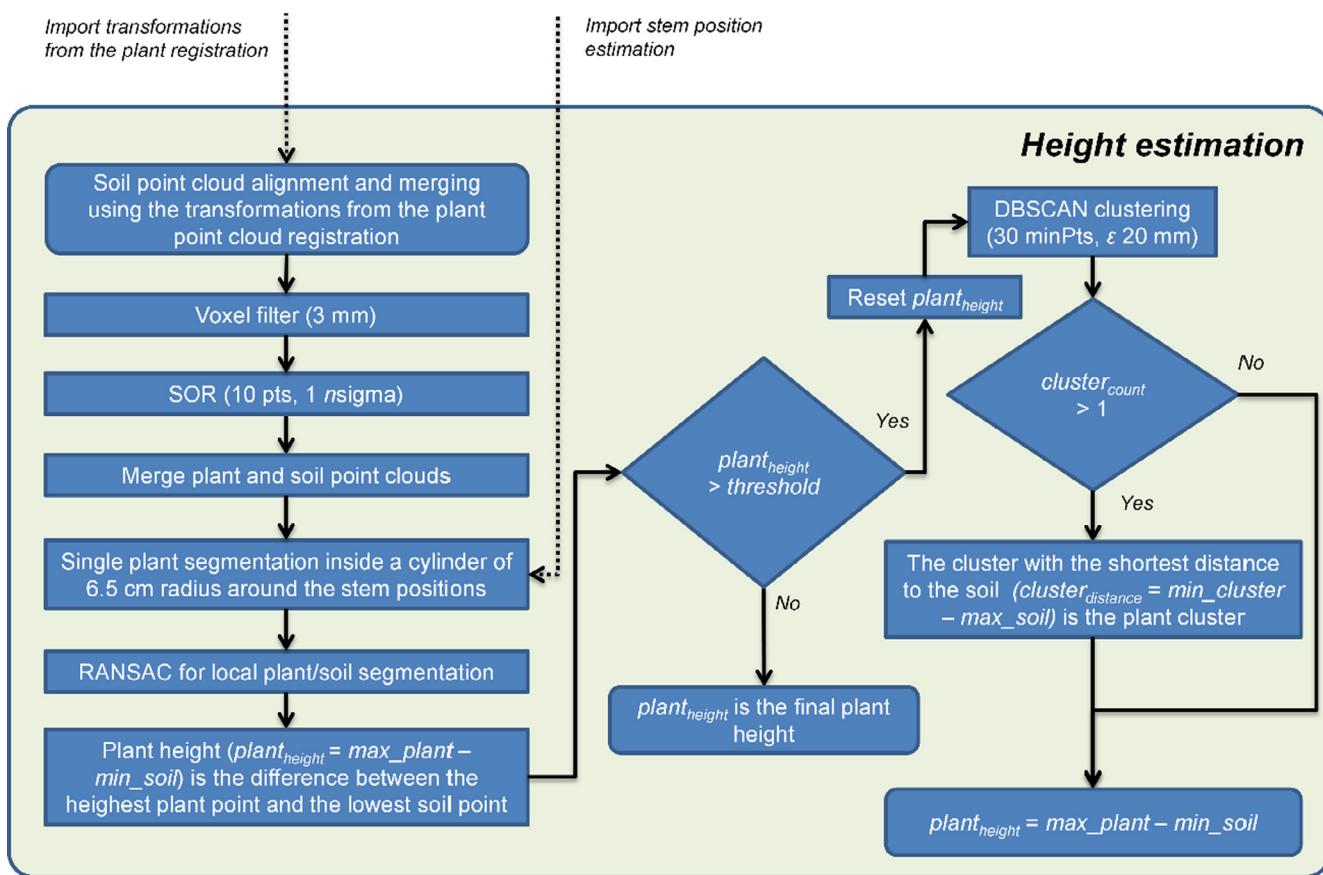


Fig. 6. Plant height estimation flowchart showing the pipeline process starting with the soil point cloud alignment, using the imported transformations for plant registration, single plant segmentation, plant clustering and height estimation.

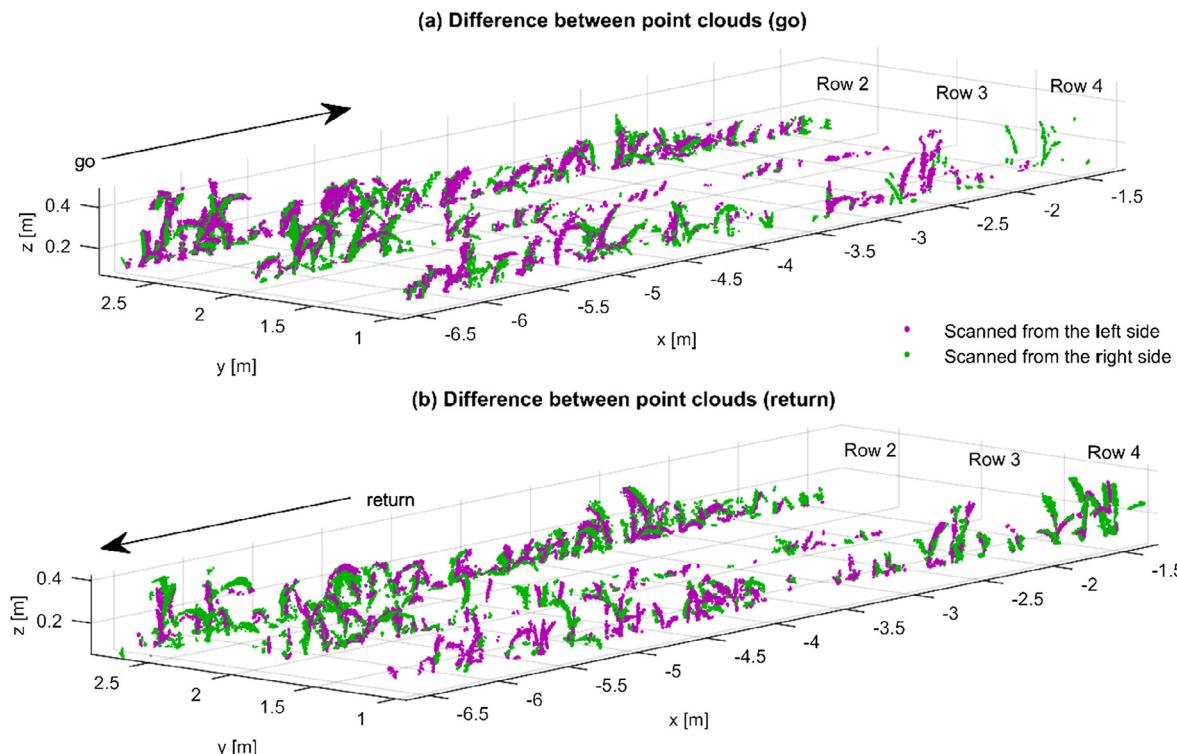


Fig. 7. Registration and alignment of point cloud pairs after the ICP algorithm. The point clouds in magenta were reconstructed when the robotic platform drove through the path at the left side of each crop row, while the ones in green were reconstructed when it drove at the right side. The driving direction was (a) go and (b) return.

Table 1

C2C results of the point cloud overlapping. The four point cloud merge- (go left + go right) + (return left + return right)- was used for the stem detection and plant height estimation of every crop row.

Crop row	Direction	Mean distance error [mm]	Standard deviation [mm]
2	go left + go right	18	23
	return left + return right	20	18
	(go left + go right) + (return left + return right)	13	12
3	go left + go right	11	10
	return left + return right	23	27
	(go left + go right) + (return left + return right)	11	10
4	go left + go right	30	44
	return left + return right	22	24
	(go left + go right) + (return left + return right)	26	27
Average		19	21

2.3.3. Plant height estimation methodology

If the plant stem position is estimated with good precision, the plant height can also be precisely estimated by calculating the difference between the maximum height, in the z axis, of the plant point cloud and the minimum height of the soil point cloud. The reason for using the minimum height is that here the miscellaneous errors are more evident since there were many overlapping soil points, if the maximum height of the soil were considered, the plant heights would be underestimated.

This approach applies when the maize plants do not overlap each other (V1-V2 growth stage), however, when the leaves are already long enough to invade the space of the neighbouring plants, the previously mentioned approach do not apply any more. Therefore, a flexible new approach was considered for plant height estimation and it is better described in Fig. 6.

The methodology for plant height estimation started by importing the rigid transformations used to register the plant point clouds in the stem estimation process. Those same transformations were necessary to align and merge the soil point clouds. It is true that the soil point clouds

were not necessary for the stem estimation process, however, they were indispensable for the plant height estimation since they provided the reference point for calculating the distance to the point with the maximum height. Additionally, it is important to emphasise that when the plants were initially segmented, using the RANSAC algorithm, in the plant stem estimation pipeline, some plant point were lost, as stated by Garrido et al. (2015), and classified as soil points (which was preferable as soil points classified as plant points), particularly the near soil surface stem points. Those points were recovered here in the plant and soil merging process. After the merging, it was required to perform a filtering, due to duplicate and flying points, using a voxel grid ($3\text{ mm} \times 3\text{ mm} \times 3\text{ mm}$) and SOR ($10\text{ pts. } 1\text{ nsigma}$) filters, respectively. At this point, with the cleaned point cloud from the final merging, the single plant segmentation process could be performed by importing the estimated plant stem positions.

In the research of Reiser et al. (2017), they developed a plant detection algorithm that was able to cluster entire plants at different growth stages, even when the leaves invaded the neighbours' space or overlapped, between V1 and V6 growth stages. However, since most of the plants in this research were between V1 and V3, with few overlapping leaves, another approach was taken by segmenting the individual plants within the boundaries of a cylinder of 65 mm radius, which was the midpoint of the plant spacing (130 mm). The stem positions were the centres of the projected cylinders (parallel to the z axis) and every point outside them was considered an outlier. Also, the Density-based spatial clustering of applications with noise (DBSCAN) algorithm (Ester et al., 1996) was used to solve the problem of leaves invading the neighbours' space.

After all the individual plants were segmented, the resulting point cloud was, once again, a combination of plants and soil points. With this point cloud we can calculate the height of most of the pants, however, if a neighbouring leave or leaves invaded the space of the targeted plant, the calculated height could be largely overestimated. Therefore, a further plant and leaf segmentation process was required, before that, firstly a local RANSAC algorithm was used to perform a local plant and soil segmentation.

In order to solve the problem of having the segmented target plant

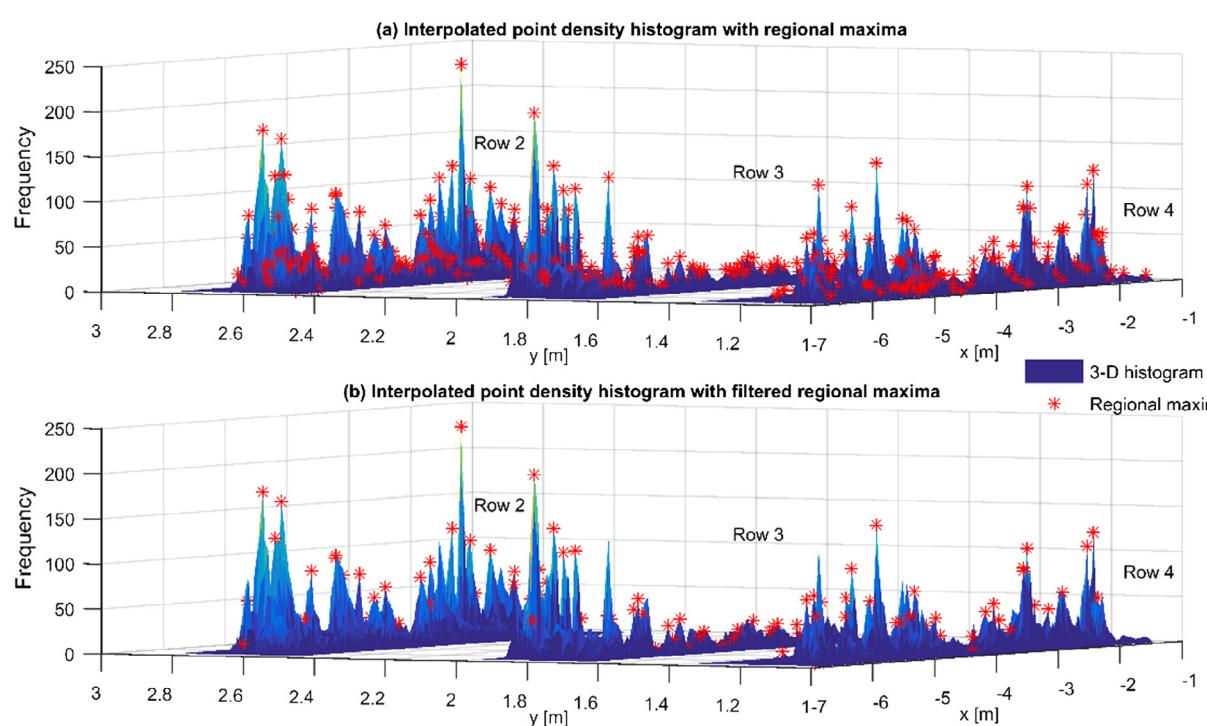


Fig. 8. Interpolated point density histogram with regional maxima (*) (a) before and (b) after the radius filtering.

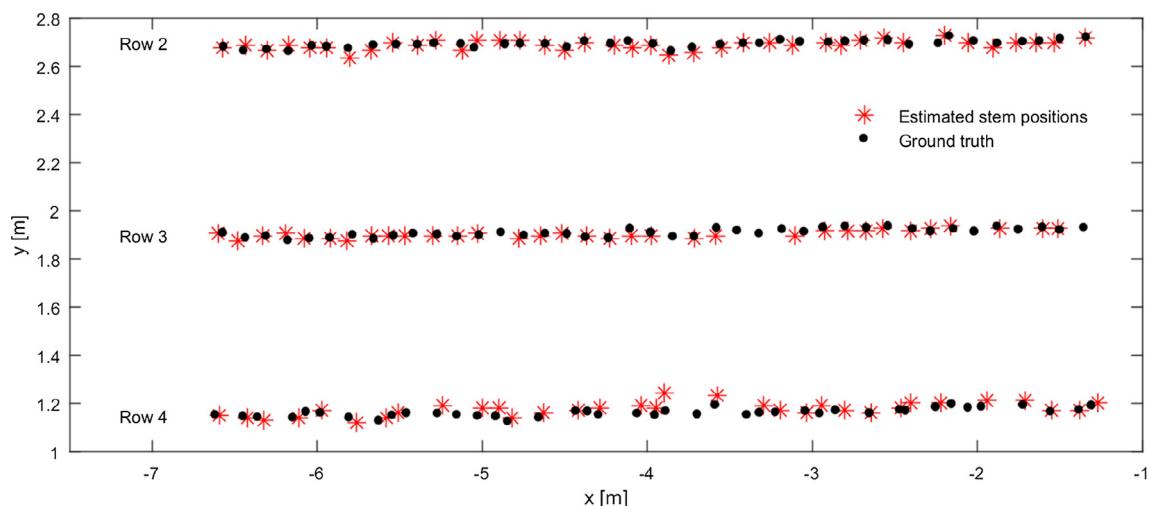


Fig. 9. Plant stem position estimation assessment. The ground truth are the seedling positions measured with the robotic total station, with sub-centimetre accuracy, after the seeding.

Table 2
Plant stem position estimation.

Crop row	Correctly detected Stems [no.]	Not detected stems [no.]	False positives [no.]	False negatives [no.]	Mean error [mm]	Standard deviation [mm]	RMSE [mm]
2	38	2	0	0	25	13	29
3	35	0	0	8	22	14	28
4	33	5	1	3	34	15	47
Average	86.2%	4.8%	0.8%	8.9%	27	14	34

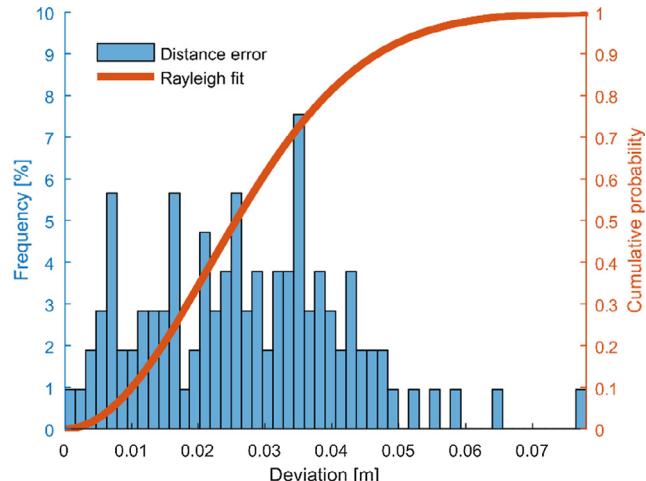


Fig. 10. Distance error histogram showing the deviation of the estimated heights in all the analysed maize rows. The cumulative probability was best fitted by a Rayleigh distribution.

together with an invading leaf or leaves from the neighbouring plants, a clustering algorithm was required to avoid using the invading leaves for the local plant height calculation, instead, the highest z value of the plant cluster plant must be obtained. Algorithms, such as the k -means, that require to specify a priori the number of clusters in the data subset would have limited success, since the number of points and location of the foreign leaf or leaves in the cylinder space could make the previously mentioned algorithm to fail, since it is difficult to predict the number of invading leaves and their location. The DBSCAN was selected because of its flexibility with the clustering of arbitrary shapes and robustness with noisy point clouds. The DBSCAN algorithm requires, as input parameters, the minimum number of points ($minPts$) in the neighbourhood (set to 30 points) and the maximum distance

between points (ε) in a cluster (set to 20 mm). The cluster with the shortest distance, from the lowest point of the cluster ($min_cluster$) to the highest point of the soil (max_soil), was considered as the plant cluster. Finally, the plant height was calculated with the difference between the maximum height of the plant cluster (max_plant) and the minimum height of the soil point cloud (min_soil).

2.3.4. Plant height profile estimation methodology

Very often for precision agriculture, it is very useful to have a plant height profile, as precise as possible, for performing an application that needs to maintain a constant offset from the plant height, like a sprayer's boom levelling system. The calculation of the plant profile of the merged point cloud followed the next steps:

- (1) Rasterize the plant point clouds using a 390 mm grid, three times the inter-row plant separation of 130 mm, projecting the maximum height of every cell in the z direction.
- (2) Rasterize the soil point cloud using a 3 mm grid, thus maintaining the same point density, and projecting the minimum height of every cell in the z direction.
- (3) Generate a mesh, or 2-D soil profile model, from the soil point cloud using a thinner grid of 50 mm to reduce the roughness and sensitivity to noise.
- (4) Compute the cloud to mesh distance (nearest neighbour) of every rasterized plant point to the meshed soil.

3. Results and discussion

3.1. Plant stem position estimation

The merging of the four point clouds from different 3-D perspective views provided, in most of the cases, complementary information for the construction of a better plant representation as seen in Fig. 7. It can be seen the difference between the point cloud pairs, where the leaf

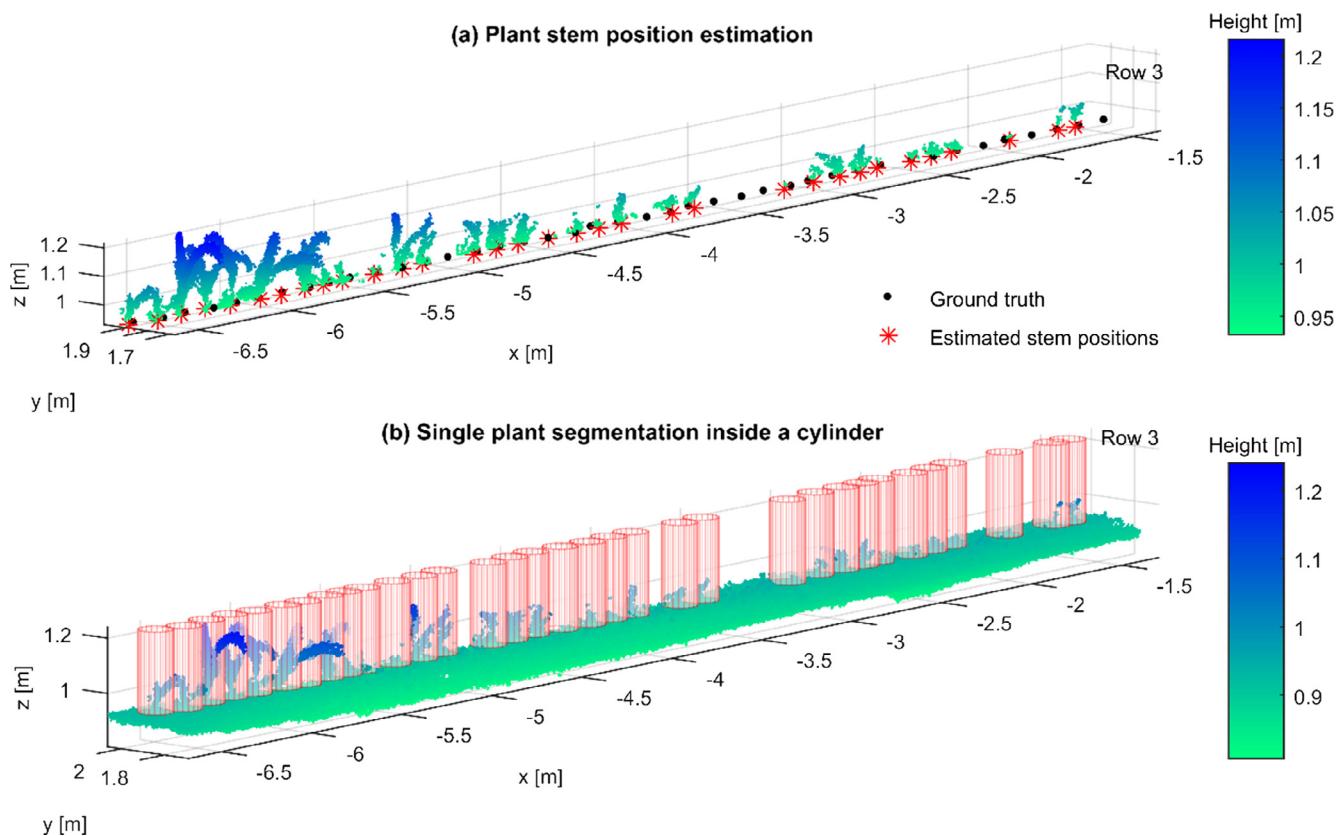


Fig. 11. Point cloud with (a) plant stem position estimation compared with the ground truth. The stem position estimations were used to project cylinders around them (b), for single plant segmentation. Every plant point outside its cylinder, was considered as an outlier. The z coordinate is related with the real plant height.

segments of some plants were missing in one point cloud, while visible in the other. Moreover, there were entire plants that were missing in one point cloud, while visible in the other. For example, in Fig. 7a, it can be seen that the last plants at row 4 ($x = -1.5\text{ m}$, $y = 1.25\text{ m}$) were not visible in the scan from the left side and barely visible for the right side while going, meanwhile, in Fig. 7b, the same plants were very well defined while returning from both sides. Also, in Fig. 7a, in row 2 ($x = -1.4\text{ m}$, $y = 2.7\text{ m}$) the last plant was visible in the scan from the right side, but not in the one from the left side.

The merging of several point clouds was necessary due the inherent occlusion in maize crop rows since some parts were out of the reach of the perspective view of the ToF camera, but as previously mentioned, the disadvantage was that miscellaneous errors made the plant thicker as it is in reality. Table 1 shows the quantification of the overlapping differences, where the worse one occurred at row 4 with a mean distance error of 30 mm while the best was 11 mm in row 3. However, the average mean and standard deviation of 19 mm and 21 mm respectively, represent a good quality in the overlapping. It must be clarified that the benefit of a better representation of the crop row (in the form of a merged point cloud) outweighs the cost of a small reduction in overlapping quality when the individual point clouds were generated with high precision (as it is the case in this research), but if the merged point cloud lacks precision, such as drifting or wrong registration, the cost-benefit would be low. Also, if the overlapping percentage between subsequent 3-D images during the registration and stitching process is high, more information is obtained and thus less crop row point clouds would be required. In this research the overlapping was not high, particularly at the end of the rows, therefore, four crop row point clouds were required for a better representation.

The interpolated point density histogram shown in Fig. 8a was the result of the computation of the regional maxima of the bivariate point density histogram. The number of detected regional maxima points was

very sensitive, aside from the bin size, to the roughness of the point cloud. In other words, if the point cloud that represents a leaf was either noisy, blurred or incomplete, it would produce more regional maxima points since the local bin count is more uneven. It was desirable to have less regional maxima points since it increases the accuracy of the radius filtering. In the case of the bin size, in this research a square bin size of 1 cm^2 was used, but if a bigger bin size was used, it would have facilitated the stem detection, due to a reduced number of regional maxima, but at the price of a degradation in precision.

Since there were more regional maxima than number of stems, a radius filter was implemented to remove the outliers. After filtering, most the regional maxima remained at the peaks, while the low-lying points (most of them outliers) at the foothills of the high peaks were removed. The result after applying the radius filter is shown in Fig. 8b.

As it can be noticed in Fig. 9, the mean error closely correlated with the standard deviation of the seeding; the more unevenly the seeding was, the larger the stem position estimation error, therefore, it is not coincidence that the row with the worse standard deviation during seeding (row 4) had the worst mean error of 34 mm.

The radius filter algorithm made an initial estimation, to detect a new plant position, expecting to be located 130 mm apart from the reference plant in the x direction. Therefore, the theoretical distance between plants is used as an input for the radius filter. The plant stem position was estimated with an average mean error and standard deviation and RMSE equal to 27 mm, 14 mm and 34 mm respectively (see Table 2). The worst mean error, as expected, was in row 4 due to the high standard deviation in the seeding, which is visually evident in Fig. 8 where the ground truth plot of row 4 is more scattered compared to row 3. Also, the high standard deviation in the seeding of row 4 put the algorithm into more difficulties, as seen in the less precision of the estimated stem positions compared to row 3 (see Fig. 9). Additionally, in row 3 there were 8 false negatives that were eliminated either by the

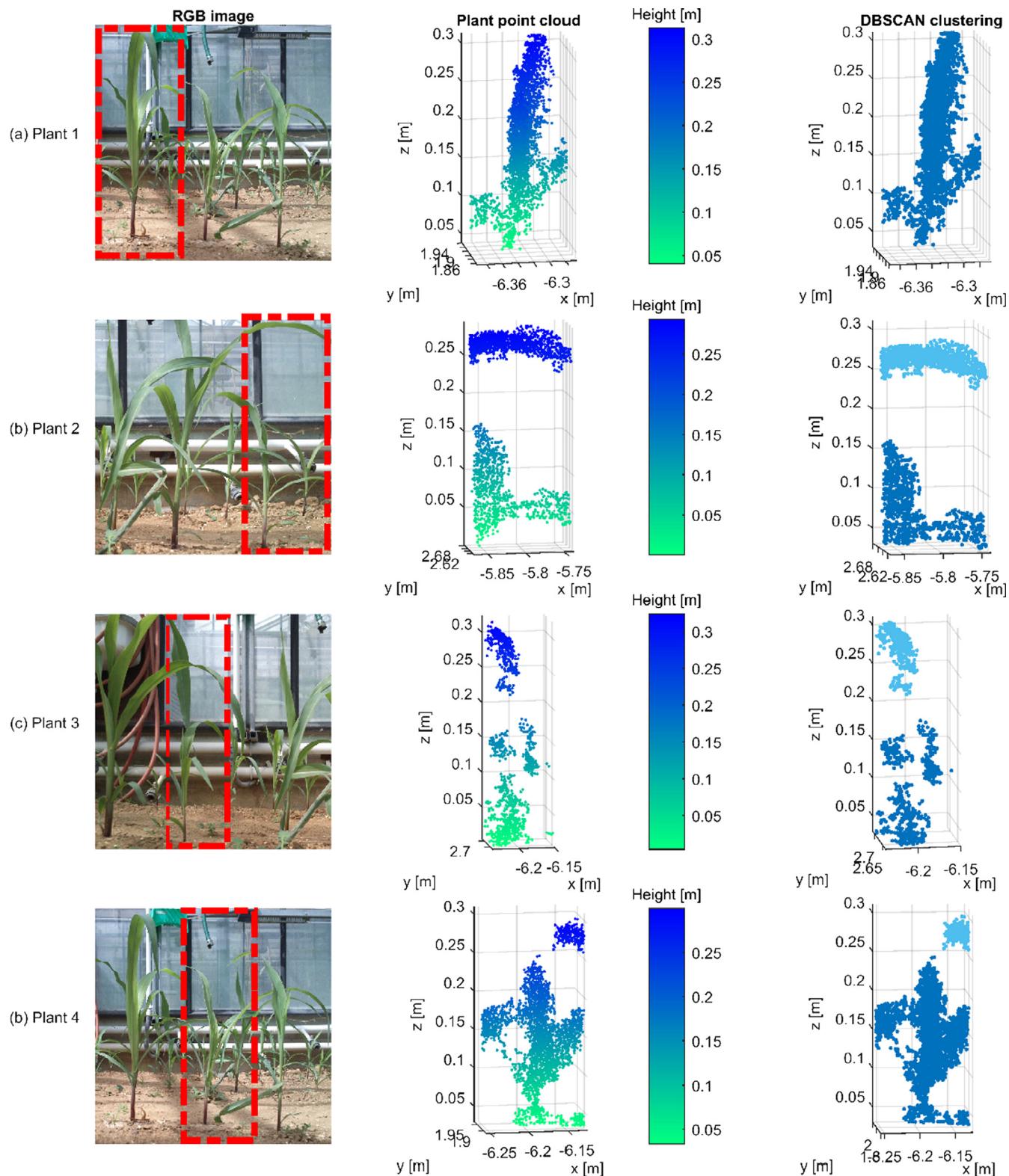


Fig. 12. DBSCAN clustering algorithm in 4 different plants (a–d). The dark blue clusters represent the plant and the bright blue the invading leaf.

RANSAC or by the filtering algorithms. Without considering the false negatives, the algorithm was able to correctly detect approximately 95% of all the stem positions.

This process was done iterative until all plant positions had been processed. Fig. 10 shows the distance error histogram with the 95th percentile equal to 4.8 mm. The distance error was best fitted by a Rayleigh distribution with a scale parameter of 21 mm. It was noticed

that with the 3-D imaging system used in the research, it was difficult to detect plants with heights lower than 130 mm or with stem diameter lower than 1.5 mm.

3.2. Plant height estimation

After the stem position was estimated, as seen in Fig. 11a, the next

Table 3
Plant height estimation.

Crop row	Correctly detected stem height [no.]	Mean error [mm]	Standard deviation [mm]	RMSE [mm]
2	38	25	37	48
3	35	20	33	40
4	33	47	35	62
Average	86.2%	30	35	50

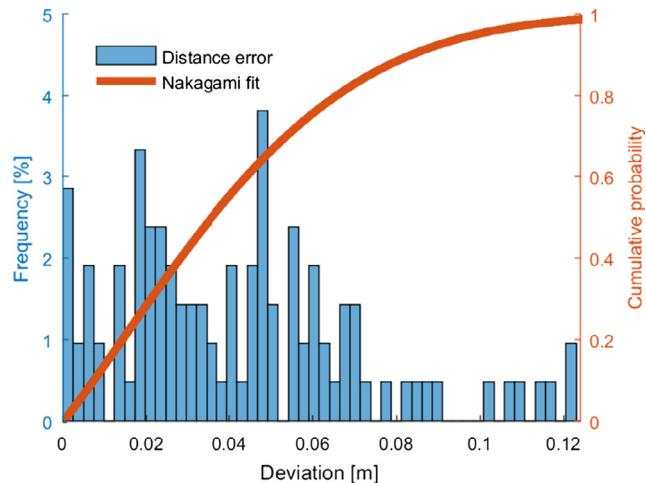


Fig. 13. Distance error histogram shows the deviation of the estimated heights of all the analysed plants. The cumulative probability was best fitted by a Nakagami distribution.

Table 4
Plant height estimation.

Crop row	Rasterized plant points [no.]	Maize plants [no.]	Overall mean error [mm]
2	14	41	6
3	13	41	8
4	14	41	12
Average			8.7

step was to use this information for individual plant segmentation. In Fig. 11b, the 65 mm radius cylinders are depicted on the final merged point cloud (four plant point clouds merging together with four soil point clouds merging) that were used as boundaries for the individual plant segmentation. Some plant points of long leaves fall outside the cylinder as seen in Fig. 11b (two long leaves between $x = -5.5$ m and $x = -6.5$ m), in this cases, the plant height would be underestimated because the real maximum height lies on the outlier. Therefore, this approach could start with considerable inaccuracies with plants from V4 growth stage onwards, since a considerable part of the leaves would lay out of the cylinder. However, as previously said, since plants of the datasets presented in this research are mostly between V1 and V3 growth stages, the overall height estimation methodology was not affected by those few underestimations.

After the plants were segmented, a RANSAC algorithm was implemented to separate the plant points from the soil points in order to discriminate, within the plant points, between the ones that were part of the local plant and the ones belonging to neighbouring leaves invading the local plant space. An important part of the proposed methodology was the implementation of the DBSCAN clustering, without it, large height underestimations would decrease the accuracy of the height estimation and therefore the viability of the proposed

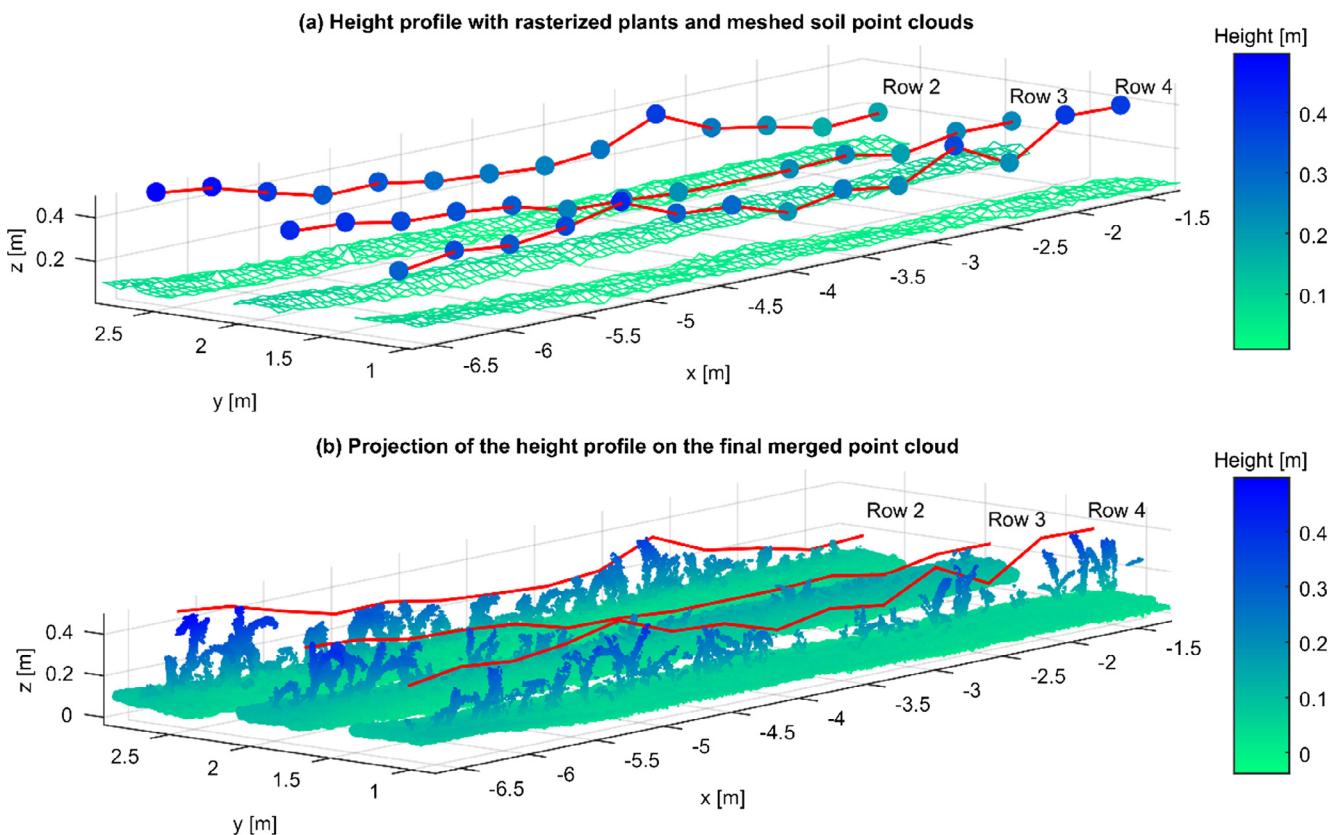


Fig. 14. (a) Height profile representation with rasterized plant points and meshed soil point clouds. (b) The merged point cloud closely agrees with the height profile (red line).

methodology. In Fig. 12a it can be seen a single plant clustering, which applied to most of the plants, because as previously said: most of the plants were inside the boundaries of the cylinder. Fig. 12b–d shows the successful discrimination between the target plant and the invading leaves from the neighbouring plants. In Fig. 12c, it is shown that Plant 3 (height = 200 mm, stem diameter = 3.5 mm) is between two bigger plants that have occluded it during the 3-D data acquisition, therefore, the points in the plant point cloud were sparse, however, the DBSCAN algorithm was able to cluster them together and differentiate them from the incoming leaf from the neighbour plant. Although, the DBSCAN algorithm was very effective, it could not correctly cluster plants with high leaf overlapping, additionally, there were several cases of folded leaves, that were too thin to be detected by the ToF camera, but their end points (apex or tip) were the highest point of the plant and were taken as a reference for the manual height measurement.

To validate the plant height profile the mean height of every row was analysed. Table 3 shows that the average mean error and standard deviation and RMSE were 30 mm, 35 mm and 50 mm, respectively. The worst mean error was in row 4, which partly was affected by the worsening of the stem estimation due to the high variability in the plant positions. Aside from that, other sources of errors were the lack of 3-D data on small plants and folded leaves that produced underestimated heights. Fig. 13 shows the distance error histogram with the 95th percentile equal 0.1 m and its best fit (Nakagami distribution), with the shape and scale parameters equals to 0.53 and 0.0026, respectively. The main source of inaccuracies for the height estimation were mainly the lack of 3-D data of the plant, therefore, although the plants higher than 130 mm were detected, the ones smaller than 200 mm or with stem diameter smaller than 30 mm, were more an abstract plant point cloud with not a clear and objective 3-D plant representation. Therefore, neither plant phenotyping nor a very precise height estimation could be performed, signalling the limitations of the 3-D ToF camera.

3.3. Plant height profile estimation

Results of the plant height profile (see Fig. 14) not just visually correlate with the final reconstructed crop row, but also numerically. The rasterized points of the plant height profile could not be compared with single plant heights, since it was an average of several plant heights in a $0.39 \text{ m} \times 0.39 \text{ m}$ cell, therefore, the assessment was done by comparing the mean height of the rasterized points with the mean height of the ground truth per row. Table 4 shows very precise results for the height profiles, reaching 6 mm error in row 2 and 8.7 mm in the average of the three rows.

4. Conclusions

The results demonstrated that the ToF camera-based 3-D imaging system was able to estimate the stem position with accuracies of 27 mm and 14 mm standard deviation. It also provided meaningful information about the plant height profile with an average overall mean error of 8.7 mm. Since the maize plants considered in this research were highly heterogeneous in height, some of them had folded leaves and were planted with standard deviations that emulate the real performance of a seeder; it can be said that the experimental maize setup was a difficult scenario, since maize heights are more homogeneous in a standard agricultural scenario. Ground-based 3-D image acquisition provided important data about the plant stem, and although the acquisition is slow (between 0.02 and 0.04 ms^{-1}) compared with an unmanned aerial system, the latter would not be able to obtain plant stem data. Another approach would have been to place the camera in a side-view position in order to obtain more data about the plant stem. Finally, having a 3-D reconstruction of the maize plants using a cost-effective sensor, mounted on a small electric-motor-driven robotic platform, means that the cost (either economic, energetic or time) of generating every point in the point clouds is greatly reduced compared with

previous researches. Further research needs to be done with the use of clustering algorithms for single plant segmentation (e.g. Euclidian, min-cut based, region growing segmentation) so that parts of long leaves are not cut off from the plant point cloud, and that plants at a latter growth stage can be also analyzed. Also, non-rigid registration and alignment needs to be explored to reduce the effect of the miscellaneous errors that reduce the overlapping between point cloud pairs. Additionally, the LAI with the data generated with this ToF camera needs to be analyzed at different growth stages to assess its correlation since it is a very important plant parameter that needs to be taken into consideration.

Acknowledgement

The project was conducted at the Max-Eyth Endowed Chair (Instrumentation & Test Engineering) at Hohenheim University (Stuttgart, Germany), which is partly grant funded by the Deutsche Landwirtschafts-Gesellschaft e.V. (DLG). The authors gratefully acknowledge Hiroshi Okamoto and Marlowe Edgar C. Burce for contributing with helpful comments on this research paper. Also, the authors would like to thank the Deutscher Akademischer Austauschdienst (DAAD) and the Mexican Council of Science and Technology (CONACYT) for providing a scholarship for the first author.

Conflict of interest

The authors declare no conflict of interest.

References

- Andújar, D., Dorado, J., Fernández-Quintanilla, C., Ribeiro, A., 2016. An approach to the use of depth cameras for weed volume estimation. Sensors (Switzerland) 16, 1–11. <https://doi.org/10.3390/s16070972>.
- Besl, P., McKay, N., 1992. A method for registration of 3-D shapes. IEEE Trans. Pattern Anal. Mach. Intell. <https://doi.org/10.1109/34.121791>.
- Blackmore, B.S., Griepentrog, H.W., Fountas, S., 2006. Autonomous systems for european agriculture conference: automation technology for off-road equipment. In: Automation Technology for off-Road Equipment. Bonn, Germany.
- Chaiivivatrakul, S., Tang, L., Dailey, M.N., Nakarmi, A.D., 2014. Automatic morphological trait characterization for corn plants via 3D holographic reconstruction. Comput. Electron. Agric. 109, 109–123. <https://doi.org/10.1016/j.compag.2014.09.005>.
- EDF R&D, T.P., 2011. CloudCompare (version 2.9.1) [GPL software] [WWW Document].
- Ester, M., Kriegel, H.-P., Sander, J., Xu, X., 1996. A density-based algorithm for discovering clusters in large spatial databases with noise. In: Simoudis, E., Han, J., Fayyad, U.M. (Eds.), Proceedings of the Second International Conference on Knowledge Discovery and Data Mining. AAAI Press, pp. 226–231.
- Fischler, M.A., Bolles, R.C., 1981. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Commun. ACM 24, 381–395. <https://doi.org/10.1145/358669.358692>.
- Garrido, M., Paraforos, D.S., Reiser, D., Arellano, M.V., Griepentrog, H.W., Valero, C., 2015. 3D maize plant reconstruction based on georeferenced overlapping lidar point clouds. Remote Sens. 7, 17077–17096. <https://doi.org/10.3390/rs71215870>.
- Gebbers, R., Adamchuk, V.I., 2010. Precision agriculture and food security. Science 327, 828–831.
- Girardeau-Montaut, D., Roux, M., Marc, R., Thibault, G., 2005. Change detection on points cloud data acquired with a ground laser scanner. IAPRSSIS 39.
- Hämmerle, M., Höfle, B., 2016. Direct derivation of maize plant and crop height from low-cost time-of-flight camera measurements. Plant Methods 12, 1–13. <https://doi.org/10.1186/s13007-016-0150-6>.
- Kahaner, D., Moler, C., Nash, S., 1989. Numerical methods and software. Prentice-Hall Inc, Upper Saddle River, NJ, USA.
- Lu, H., Tang, L., Whitham, S.A., 2015. Development of an automatic maize seedling phenotyping platform using 3D vision and industrial robot arm. In: ASABE Annual International Meeting. Louisiana, USA.
- Lu, H., Tang, L., Whitham, S.A., Mei, Y., 2017. A robotic platform for corn seedling morphological traits characterization. Sensors (Switzerland) 17, 1–17. <https://doi.org/10.3390/s17092082>.
- Nakarmi, A.D., Tang, L., 2012. Automatic inter-plant spacing sensing at early growth stages using a 3D vision sensor. Comput. Electron. Agric. 82, 23–31. <https://doi.org/10.1016/j.compag.2011.12.011>.
- Reiser, D., Garrido-Izard, M., Vázquez-Arellano, M., Paraforos, D.S., Griepentrog, H.W., 2015. Crop row detection in maize for developing navigation algorithms under changing plant growth stages. In: Robot 2015. Second Iberian Robotics Conference. Lisbon, Portugal, pp. 371–382.
- Reiser, D., Vázquez-Arellano, M., Garrido-Izard, M., Paraforos, D.S., Sharipov, G., Griepentrog, H.W., 2017. Clustering of Laser Scanner Perception Points of Maize

- Plants. Adv. Animal Biosci.: Precision Agric. (ECPA). 204–209. <https://doi.org/10.1017/S204047001700111X>.
- Reiser, D., Vázquez-Arellano, M., Paraforos, D.S., Garrido-Izard, M., Griepentrog, H.W., 2018. Iterative individual plant clustering in maize with assembled 2D LiDAR data. Comput. Ind. 99, 42–52. <https://doi.org/10.1016/j.compind.2018.03.023>.
- Ribeiro, A., Bengoechea-Guevara, J.M., Conesa-Muñoz, J., Nuñez, N., Cantuña, K., Andújar, D., 2017. 3D monitoring of woody crops using an unmanned ground vehicle. Adv. Anim. Biosci. 8, 210–215. <https://doi.org/10.1017/S2040470017001200>.
- Ritchie, S., Hanway, J., Benson, G., 1992. How a Corn Plant Develops. Ames, USA.
- Rusu, R.B., Cousins, S., 2011. 3D is here: point cloud library. IEEE Int. Conf. Robot. Autom. 1 – 4. 10.1109/ICRA.2011.5980567.
- Steckel, T., 2017. Entwicklung einer kontextbasierten Systemarchitektur zur Verbesserung des kooperativen Einsatzes mobiler Arbeitsmaschinen. University of Hohenheim.
- Steder, B., Konolige, K., 2011. Point Feature Extraction on 3D Range Scans Taking into Account Object Boundaries. In: 2011 IEEE International Conference on Robotics and Automation. Shanghai, China, pp. 2601–2608. 10.1109/ICRA.2011.5980187.
- Vázquez-Arellano, M., Griepentrog, H.W., Reiser, D., Paraforos, D.S., 2016a. 3-D imaging systems for agricultural applications — a review. Sensors 16, 618. <https://doi.org/10.3390/s16050618>.
- Vázquez-Arellano, M., Reiser, D., Garrido, M., Griepentrog, H.W., 2016b. Reconstruction of geo-referenced maize plants using a consumer time-of-flight camera in different agricultural environments. In: Ruckelshausen, A., Meyer-Aurich, A., Rath, T., Recke, G., Theuvsen, B. (Eds.), Intelligent Systems - Stand Der Technik Und Neue Möglichkeiten. Gesellschaft für Informatik e.V. (GI), Osnabrück, Germany, pp. 213–216.
- Vázquez-Arellano, M., Reiser, D., Paraforos, D.S., Garrido-Izard, M., Burce, M.E.C., Griepentrog, H.W., 2018. 3-D reconstruction of maize plants using a time-of-flight camera. Comput. Electron. Agric. 145, 235–247. <https://doi.org/10.1016/j.compag.2018.01.002>.