Original papers

3-D reconstruction of maize plants using a time-of-flight camera

Manuel Vázquez-Arellano[a,*], David Reiser[a], Dimitris S. Paraforos[a], Miguel Garrido-Izard[b],
Marlowe Edgar C. Burce[c], 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*
[c] *Department of Computer Engineering, University of San Carlos Talamban Campus, Nasipit, Talamban, Cebu 6000, Philippines*

ARTICLE INFO

Keywords:
3-D sensors
Kinect v2
Time-of-flight
Agricultural robotics
Precision farming

ABSTRACT

Point cloud rigid registration and stitching for plants with complex architecture is a challenging task, however, it is an important process to take advantage of the full potential of 3-D cameras for plant phenotyping and agricultural automation for characterizing production environments in agriculture. A methodology for three-dimensional (3-D) reconstruction of maize crop rows was proposed in this research, using high resolution 3-D images that were mapped into the colour images using state-of-the art software. The point cloud registration methodology was based on the Iterative Closest Point (ICP) algorithm. The incoming point cloud was previously filtered using the Random Sample Consensus (RANSAC) algorithm, by reducing the number of soil points until a threshold value was reached. This threshold value was calculated based on the approximate number of plant points in a single 3-D image. After registration and stitching of the crop rows, a plant/soil segmentation process was done relying again on the RANSAC algorithm. A quantitative comparison showed that the number of points obtained with a time-of-flight (TOF) camera, compared with the ones from two light detection and ranging (LIDARs) from a previous research, was roughly 23 times larger. Finally, the reconstruction was validated by comparing the seedling positions as ground truth and the point cloud clusters, obtained using the *k*-means clustering, that represent the plant stem positions. The resulted maize positions from the proposed methodology closely agreed with the ground truth with an average mean and standard deviation of 3.4 cm and ± 1.3 cm, respectively.

1. Introduction

Crop phenotyping is a prerequisite to enable robots perform agricultural tasks, evaluate crop status for farm management and relate genotypes to phenotypes for crop breeding among others (Vázquez-Arellano et al., 2016b); yet it remains a bottleneck (Dhondt et al., 2013) due to the time-consuming measuring methodologies and systems. Additionally, precision and organic agriculture are increasing around the world, and the many concerns about how we produce our food, particularly the chemical inputs that are not just deteriorating the environment at the time of an uncertain climate change, but also affecting human health. Moreover, if an automatic or autonomous agricultural application is considered (Reiser et al., 2017), it is important to acquire sensor data that is accurately and precisely geo-referenced in order to know the position in space of every plant and, if possible, every plant element. Although 2-D imaging can be used for estimating some plant parameters such as crop growth (Kataoka et al., 2003), they are

technically limited for obtaining actual phenotypic traits. In the other hand, as Vázquez-Arellano et al. (2016a) concluded, 3-D imaging sensors are able to provide such information at the cost of handling higher data densities and thus higher computing power. Fortunately, off-the-shelf 3-D imaging sensors have reached sufficient technical maturity to handle the task of measuring plant properties in a cost-effective manner, while parallel computing can accelerate the time-to-solution in order to alleviate the time-consuming computation of dense point clouds; these factors make 3-D imaging in agriculture more appealing and accessible.

3-D imaging sensors such as the Microsoft Kinect v2 (Microsoft, Redmond, WA, USA) has awakened great interest among researchers in the field of agriculture due to its relative robustness against sunlight. Since time-of-flight (TOF) cameras were previously expensive and with limited pixel resolution, publications using these type of sensors were not very common in agricultural research. However, several applications can be found in the literature such as a comparison between

---

Nomenclature

α	constant angle of inclination of the TOF camera, $\pi/4$ rad
a	distance from the TOF sensor to the target prism along the x axis, m
b	distance from the TOF sensor to the target prism along the z axis, m
c	distance from the TOF sensor to the base of the robotic platform, m
d	distance from the TOF sensor to the target prism along the y axis, m

e	distance from the target prism to the tip of the plummet, m
T_{tof}^{robot}	transformation matrix from the TOF sensor to the robotic platform coordinate system
T_{robot}^{ts}	transformation matrix from the robotic platform to the total station coordinate system
$x_{tof}\,y_{tof}\,z_{tof}$	coordinate system of the TOF sensor
$x_{robot}\,y_{robot}\,z_{robot}$	coordinate system of the robotic platform
$x_{ts}\,y_{ts}\,z_{ts}$	global coordinate system of the total station data
θ	pitch angle, rad
φ	roll angle, rad
ψ	yaw angle, rad

---

commercial TOF cameras and their suitability for agricultural applications (Klose et al., 2009; Kazmi et al., 2014). Also, an application for plant phenotyping that relies on two TOF cameras, among others, mounted on a multi-sensor platform (Busemeyer et al., 2013). Agricultural automation was also investigated using TOF cameras for a human-machine interactive system where an agricultural robot follows a human (Yin and Noguchi, 2013). Recently, the Kinect v2 was used for discriminating crop plants from weeds using their respective heights as the sole differentiation parameter (Andújar et al., 2016). Rosell-Polo et al. (2017) developed a sensor-based mobile terrestrial laser scanner using the Kinect v2 for vineyard characterization, and a real time kinematic-global navigation satellite system (RTK-GNSS) for geo-referentiation. They compared different reconstructions using a single column, partial and complete field of view (FOV) at different frame rates with a stable acquisition speed. They obtained the best results with the single column, which actually emulates the operation of a light sheet (2-D) light detection and ranging (LIDAR). However, due to problems with the juxtaposition (gaps between images), when the partial and complete FOV were considered, they acknowledged that the lack of scanning continuity could be a limiting factor for the applicability of the system under field conditions. Similarly, in another research performed by Butkiewicz (2014), based on a mobile platform that carries a Kinect v2, it was mentioned the complexity of point cloud rigid registration and stitching. One of the few characterizations of maize plants using a robotic platform was done by Weiss and Biber (2011) with a 3-D LIDAR. The image acquisition technique is comparable with the one of a TOF camera, since both rely on a light volume technique for depth measurement. Nevertheless, the maize leaves were not clearly defined in detail due to the poor 3-D image resolution (59×29 pixels) of the sensor and the vertical error of the RTK-GNSS.

For point cloud rigid registration, there are several researches in agriculture relying on the popular algorithm Iterative Closest Point (ICP) (Besl and McKay, 1992). Hoffmeister et al. (2013) considered a software that used the ICP to enhance the registration process, relying on differential GNSS for geo-referentiation, to generate crop surface models using a terrestrial 3-D LIDAR. Dong et al. (2017) reconstructed crop rows at different vegetative and temporal stages based on a three-step method: Multi-sensor Simultaneous Localization and Mapping (SLAM), data association and optimization to build a full 4-D reconstruction. For comparison, they used the ICP but the results were not satisfactory compared with their proposed approach citing as a limitation that the ICP can only compute a single rigid relative transformation for each point cloud pair, while their method can perform data association in multiple places, which is equivalent to a non-rigid transformation. Also, Mai et al. (2015) used the ICP to register apple trees from two different perspective views. Their proposed method relied on the search of key points, vectorization and filtering, and fast and precise registration. Since the point clouds of this research were not geo-referenced, a fast registration was required to approximate the two datasets using the Fast Point Feature Histograms (FPFH). Subsequently, a precise registration was performed using the ICP algorithm. To speed up the computation, they used parallel processing even though they

were dealing with single 3-D image pairs.

The aim of this research was to reconstruct maize plants, based on the ICP algorithm, using a TOF camera. A qualitative analysis was performed by comparing the generated point clouds with the colour Red-Green-Blue (RGB) image representation, and the published results of a previous research paper by Garrido et al. (2015) using the two LIDARs (light sheet) mounted on the same robotic platform during the experiments. Furthermore, a validation was done by comparing the estimated stem positions of the generated point cloud with the ground truth.

2. Materials and methods

2.1. Hardware and sensors

A robotic platform, called TALOS, is a small four-wheel autonomous robot developed at the University of Hohenheim. The characteristics of the TALOS robotic platform are explained in detail by Reiser et al. (2015). This robotic platform was used for data acquisition in a greenhouse and carried multiple imaging sensors for data acquisition, however, in this research paper only the Kinect v2 TOF camera is considered. The robotic platform was also equipped with a VN-100 Inertial Measurement Unit (IMU) (VectorNav, Dallas, USA) used to obtain the orientation of the entire acquisition system as it navigated along the paths. The IMU sensor, situated inside at the centre of the robotic platform, was particularly useful since the soil was uneven and the internal gyroscopes provided 3-D angular measurements to obtain the orientation of the whole acquisition system. An IMU sensor is indispensable for any practical use of 3-D sensors in agricultural applications to compensate the effects of rough terrain and mechanical vibrations to the 3-D imaging system. The SPS930 robotic total station (Trimble Navigation Limited, Sunnyvale, USA) was used to track the position of the vehicle by aiming at the Trimble MT900 Machine Target Prism, which was mounted on the top of the vehicle at a height of 1.07 m (Fig. 1). The positioning data was transmitted and stored in the Yuma 2 Tablet, which was connected to the robotic platform computer for data exchange (Reiser et al., 2015). The prism position, IMU and TOF camera data were time stamped with a sampling rate of 20 Hz, 50 Hz and 5 Hz; respectively, and were used to refer the static transformations to the global coordinates given by the total station frame.

The robotic platform software, for both navigation and data acquisition, was developed using the Robot Operating System (ROS Indigo), an open source middleware for robot software development running on Linux (Ubuntu 14.04), and programmed in a combination of C++ and Python programming languages. For calibration, point measurement, and importing data from the Total Station into ROS, the Trimble SCS900 Site Controller (Software Version 3.4.0) graphical interface was used; the Trimble SCS900 Site Controller was installed in the Yuma 2 Tablet. For a better understanding of the computer hardware and software used in this research, Table 1 shows a more detailed explanation.

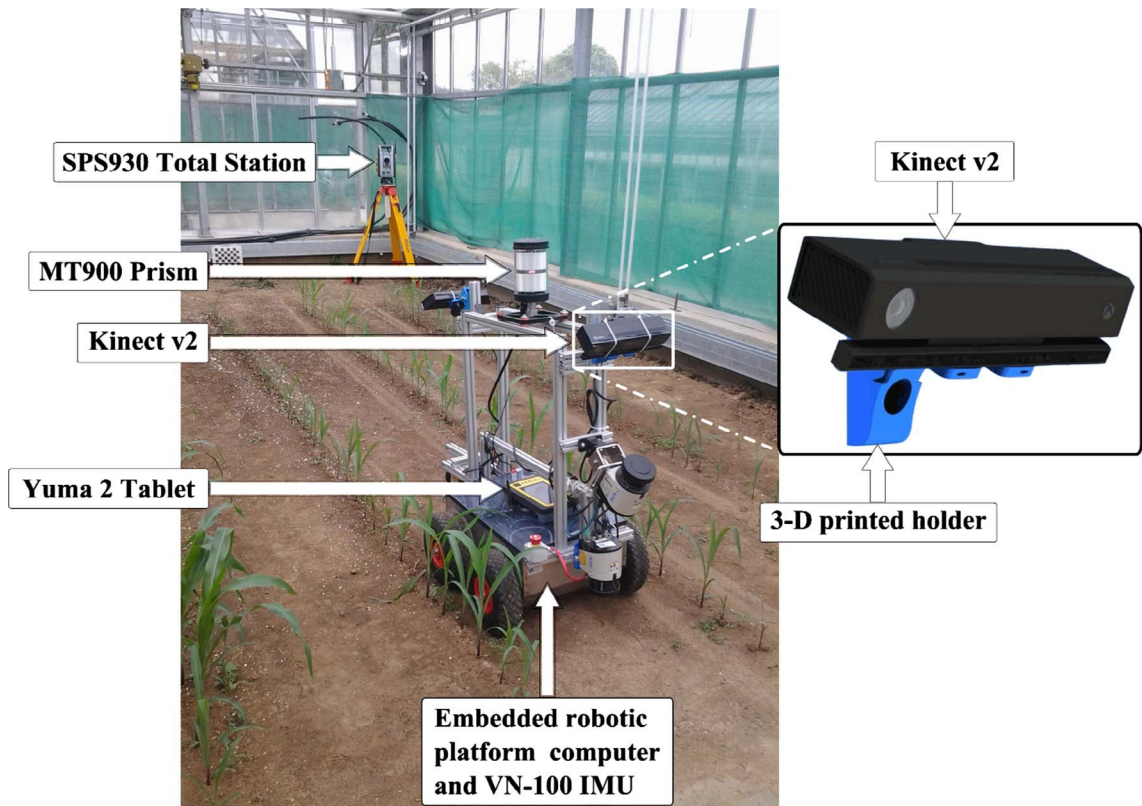The basic principle for optical depth measurement behind the

Fig. 1. Acquisition system for the experiment depicting the TALOS robotic platform carrying multiple sensors. For this research paper, the data from the Kinect v2, VN-100 IMU and the geo-referenciation system were used. The geo-referenciation system consists of the SPS930 total station, MT900 target prism and Yuma 2 Tablet.

Kinect v2 sensor is time-of-flight, specifically continuous wave modulation TOF (Lachat et al., 2015a,b). The main technical characteristics of this particular TOF camera are its high-resolution colour and depth output stream. Additionally, it has an infrared output stream that could enhance current 2-D night vision applications, or systems that rely on shadowing devices. The TOF camera had a measurement range between 0.4 and 4.5 m, and it was mounted at a height of 0.94 m with a downwards view at an angle of 45°. The nominal sampling rate of the TOF camera was 30 Hz; however, to avoid losing data by overloading the acquisition system, the frames per second had to be shortened to 5 Hz. The TOF camera outputed three image streams: depth, infrared and colour RGB; a more detailed description of the TOF camera technical specifications is shown in Table 2. In order to hold the TOF camera properly while being able to modify the angle of inclination, a holder was designed in SolidWorks and was 3-D printed (see Fig. 1).

After performing the data acquisition with the aforementioned hardware and software, the data processing was done using a high-end fanless Pokini Workstation. C++ programing in ROS was used to perform the image registration and stitching. Then, the Robotics System Toolbox™ of MATLAB R2016b was used as an interface for exporting the stitched point cloud from ROS to MATLAB for further segmentation and visual representation.

2.2. Experimental setup

The experiment was conducted in a greenhouse of 21 m^2 (width = 3.75 m, length = 5.6 m) at the Agricultural Technology Centre Augustenberg at the University of Hohenheim, Stuttgart, Germany (48°42′50.9″N 9°12′30.7″E). The average outdoor temperature during the experiments was 13.9 °C, while the greenhouse's temperature was automatically controlled ranging from 22 °C to 25 °C. The maize was irrigated on a daily basis, and was fertilized two times during the experiments without removing the weeds.

A waiting time of approximately 1 h between switching on the camera and performing the field experiments was taken to stabilize the depth measurements. Warming up the Kinect v2 is advisable since the measurements can vary almost 5 mm from the switch-on time until they start stabilizing after 30 min, when the measurements vary ± 1 mm (Lachat et al., 2015b). Also, a calibration was performed before the data acquisition in order to refer to the same global coordinates of the total station frame (Reiser et al., 2016). The data was then acquired by driving the robot through the 4 paths (5 crop rows) using a remote

Table 1
General specifications of the computer hardware, their operating systems and installed software.

Hardware	Specification	Function	Operating system	Software
Robotic platform computer	i3-Quadcore processor (3.3 GHz), 4 GB RAM and SSD hard drive	To control the robotic platform and provide connectivity to receive and store different sensor data	Ubuntu 14.04	ROS
Yuma 2 Tablet	Intel Atom COU N2600 dual-core processor (1.6 GHz), 4 GB RAM and SSD hard drive	To receive positioning data from the SPS930 via a wireless link (2.4 GHz IEEE 802.11) and export it to the robotic platform computer via RS232	Windows 7 Professional	Trimble SCS900 site Controller
Pokini workstation	Core i7 processor (3.3 GHz), 32 GB RAM, 8 GB NVIDIA Quadro M4000 graphic card and SSD hard drive	It was used exclusively for data analysis	Ubuntu 16.04	ROS and MATLAB

Table 2
General specifications of the TOF camera (Microsoft, 2017).

Feature	Specification	Value
General data	Model	Kinect v2
	Cost (cables included)	~140 €
	(length × width × height)	24.9 cm × 6.6 cm × 6.7 cm
	Weight	1.4 kg
Depth sensing	Basic principle	Time-of-flight
	Depth range	0.4 m–4.5 m
	Depth image resolution	512 × 424 pixels
	Field of view	70° × 60°
	Frame rate	30 Hz
Colour camera	Colour image resolution	1920 × 1080 pixels
	Frame rate	30 Hz (15 Hz in low light)
Active infrared	Infrared image resolution	512 × 424 pixels
	Frame rate	30 Hz
	Infrared light wavelength	~827–850 nm
Data transmission	Interface standard	USB 3.0

joystick at a constant speed of *circa* $0.05 \, \text{m s}^{-1}$ (Garrido et al., 2015). The dataset for this research was acquired on the 30th of April 2015 when the maize plants were between V2 and V4 vegetative stages (Ritchie et al., 1992). The seeding was done considering a plant distance of 0.13 m, but in order to emulate the accuracies of real seeding, different Gaussian distributions were considered by generating random distance errors, with a specific standard deviation per row, that differ from the ideal seeding. The standard deviation of the rows were generated with an excel sheet random function. For crop row 1–5, the standard deviations in the seeding were 0.019, 0.017, 0.006, 0.048 and 0.047 m, respectively. The row length was 5.2 m with 41 plants per row. The row spacing was 0.75 m and the headland approximately 1.5 m at each end of the planted maize. To obtain the ground truth, a plummet was attached to a tripod where the target prism was mounted, and for all the seedlings in the greenhouse the position was measured (as seen in Fig. 2); by moving the tripod, target prism and plummet to the position of every seedling. After all the seedling positions were measured (just one time), the target prism was unmounted from the tripod and mounted on the robotic platform in order to track the position of the 3-D imaging acquisition system during the inter-row navigation when going and returning.

2.3. Data processing

2.3.1. Software and data pre-processing

A ROS package "iai kinect2" developed by Wiedemeyer (2015) was used to obtain the point cloud data from the Kinect v2. This package was particularly useful because it provides a driver to receive data from the TOF camera through the *sensor_msgs::PointCloud2* message, which is already a 3-D image. Here, it is possible to obtain low resolution or high resolution point clouds. The "iai kinect2" does a registration and mapping of the depth image (512 × 424 pixels) into the colour RGB

image (1920 × 1080 pixels). In order to register the depth image to the higher resolution colour image, the following process was performed:

● The 3-D points were computed for each depth pixel.
● A rotation and translation of the points was performed (resulting in points as seen from the position of the colour sensor).
● A mapping between the two sensors (depth and colour) was created by using the intrinsic parameters of both of them.
● A map of each transformed depth pixel to the colour sensor intrinsic parameters was generated.

During the mapping process, an up-scaling and interpolation were performed. This is why the size of the high resolution point cloud is bigger than the one generated by the depth sensor. The depth sensor resolution image could have also been used, and with that, the size of each individual 3-D image would be a quarter of the one used in this research. However, it was decided to use the high resolution point cloud not just to investigate the limits of the available options in the state-of-the-art software and hardware, but also to allow additional research possibilities such as 2-D and 3-D fusion by using algorithms based on colour information (i.e. colour-based region growing segmentation).

In order to perform sensor fusion (data from the TOF camera, IMU and total station), first the total station data was adjusted to compensate the delay in the time stamp due to a latency of 40 ms. Then, the total station and IMU data were both interpolated to the TOF camera data time stamp, since the latter was the slowest. Although sensor fusion was performed in all the datasets, for the purpose of this research only the pose (orientation and position) of the first 3-D image was directly considered, since the rest of the incoming images were registered to this first reference point cloud. Consequently, the pose of all the subsequent point clouds was modified during the registration process.

2.3.2. Point cloud rigid transformation

In order to geo-reference every single point from the TOF camera was necessary, first, to transform the camera frame defined by $(x_{tof} y_{tof} z_{tof})$ to the robot frame, located in the target prism, defined by $(x_{robot} y_{robot} z_{robot})$; and then, transform the robot frame to the total station frame, defined by $(x_{ts} y_{ts} z_{ts})$. The previously mentioned frames are depicted in Fig. 3.

To translate all the point clouds from the TOF camera to the total station frame, the first step was to transform the TOF camera frame to the robot frame as shown in Eq. (1).

$$[x_{robot} y_{robot} z_{robot} 1]^T = [T_{tof}^{robot}] \times [x_{tof} y_{tof} z_{tof} 1]^T \tag{1}$$

where T_{tof}^{robot} is the transformation matrix from the coordinate system of the TOF camera to the coordinate system of the robotic platform. This required one rotation to match correctly the axis of the TOF camera and robot coordinate systems $Rot('yaw', -\frac{\pi}{2})$, a rotation $Rot('roll', \frac{\pi}{2} - \alpha)$ and three translations $(Trans(x, +a), Trans(z, +b)$ and $Trans(y, -d))$. The complete transformation is represented in Eq. (2).
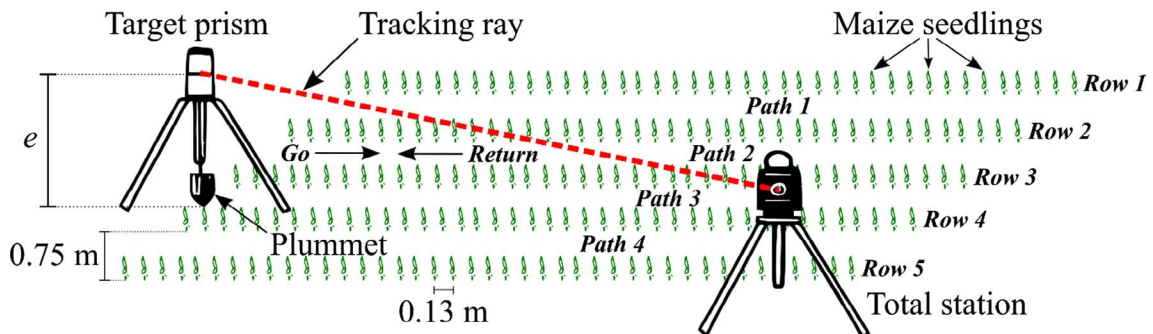

Fig. 2. Arrangement of seedling position measurement. The distance *e* is a constant value between the target prism and the tip of the plummet. The path and crop row numbering as well as the driving direction "go" and "return", and the seedling spacing are also depicted.
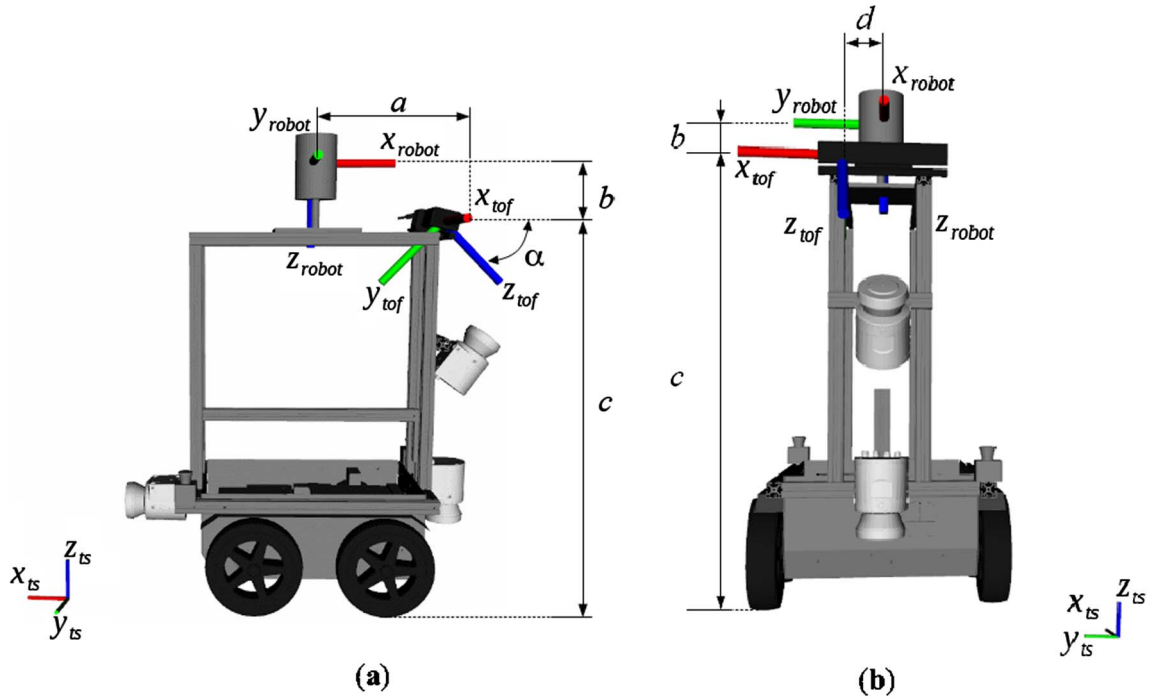
Fig. 3. Representation of the used robotic platform showing the coordinate frames of the TOF camera ($x_{tof} y_{tof} z_{tof}$), robotic platform ($x_{robot} y_{robot} z_{robot}$) and total station ($x_{ts} y_{ts} z_{ts}$) and their locations relative to each other. (a) Side view; (b) front view.

$$T_{tof}^{robot} = \underbrace{Rot\left('yaw',-\frac{\pi}{2}\right)}_{coordinate\,system\,matching} \cdot Rot\left('roll',\frac{\pi}{2}-\alpha\right) \cdot Trans(x,+a) \cdot Trans(z,+b) \cdot$$

$$Trans(y,-d) \tag{2}$$

The general matrix form of all rotations and translations are shown in Eq. (3), and by substituting them in Eqs. (2) and (5), a matrix form representation of the equations can be generated.

$$Rot('roll',\varphi) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\varphi) & -\sin(\varphi) & 0 \\ 0 & \sin(\varphi) & \cos(\varphi) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, Trans('xaxis',x) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ x & 0 & 0 & 1 \end{bmatrix},$$

$$Rot('pitch',\theta) = \begin{bmatrix} \cos(\theta) & 0 & \sin(\theta) & 0 \\ 0 & 1 & 0 & 0 \\ -\sin(\theta) & 0 & \cos(\theta) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, Trans('yaxis',y)$$

$$= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & y & 0 & 1 \end{bmatrix},$$

$$Rot('yaw',\psi) = \begin{bmatrix} \cos(\psi) & -\sin(\psi) & 0 & 0 \\ \sin(\psi) & \cos(\psi) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, Trans('zaxis',z) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & z & 1 \end{bmatrix}$$

$$\tag{3}$$

Then, the robot frame needs to be transformed to the total station frame as depicted in Eq. (4)

$$[x_{ts} y_{ts} z_{ts} 1]^T = [T_{robot}^{ts}] \times [x_{robot} y_{robot} z_{robot} 1]^T \tag{4}$$

where T_{robot}^{ts} is the transformation matrix from the coordinate system of the robot to the coordinate system of the total station. This procedure requires, firstly, coordinate system matching $Rot('pitch',-\pi)$, then, three rotations to stabilize the orientation of the robot based on the information provided by the IMU ($Rot('pitch',-\theta)$, $Rot('roll',-\varphi)$ and $Rot('yaw',-\psi)$), and finally three translations using the information provided by the total station ($Trans(x,\pm x_{ts})$, $Trans(y,-y_{ts})$ and

$Trans(z,-z_{ts})$). The complete procedure is represented in Eq. (5).

$$T_{robot}^{ts} = \underbrace{Rot('pitch',-\pi)}_{coordinate\,system\,matching} \cdot Rot('pitch',-\theta) \cdot Rot('roll',-\varphi) \cdot Rot('yaw',-\psi) \cdot$$

$$Trans(x,\pm x_{ts}) \cdot Trans(y,-y_{ts}) \cdot Trans(z,-z_{ts}) \tag{5}$$

2.3.3. Point cloud rigid registration and stitching

TOF cameras such as the Kinect v2, parallel computing, and algorithms such as the ones available in the Point Cloud Library (PCL) (Rusu and Cousins, 2011) and the Computer Vision System Toolbox™ of MATLAB R2016b (MathWorks, Natick, MA, USA) greatly facilitate the past limitations that prevented applications of these sensors in agriculture. The ICP algorithm aligns two datasets by iteratively minimizing the distances of corresponding points. In order to speed up the search process, an heuristic approach is considered by building a *k*-d tree (Bentley, 1975), which is a binary search tree where each node represents a partition of the *k*-dimensional space. Additionally, in order to balance the number of points belonging to plants and soil, the Random Sample Consensus (RANSAC) algorithm (Fischler and Bolles, 1981) was used to remove points that belonged to the soil. The purpose was to balance the plant/soil ratio, to obtain better results in the point cloud rigid registration process. The precise pose of the first point cloud was established as the reference and the subsequent point clouds were then stitched together. A prerequisite for this to work is that the data set should consist of point clouds that have been roughly pre-aligned in a common coordinate system and overlap with one another. Although the reference frame position was measured with a very accurate geo-referencing system, the methodology of point cloud rigid registration and stitching is also applicable to other less accurate geo-referencing systems. Therefore, if the ICP works well, then, the precision of the reconstruction should be defined by it, regardless of the geo-referencing system for precision agriculture purposes.

Performing point cloud rigid registration, using the ICP algorithm, where the majority of points are soil points, does not necessarily produce precisely registered plant points since the main problem when using the ICP is to determine the correct data associations. Therefore,

the aforementioned adaptive methodology was developed to increase the precision of plant point registration in order to incrementally register a series of point clouds two by two. The workflow of the proposed methodology is depicted in Fig. 4. A variant of the ICP uses the estimated local surface normal vectors and curvatures for geometric attributes to select the possible corresponding points. After computing the surface normal, a weight factor was assigned to x, y, z and the normal. Therefore, the ICP tried to minimize the distance between a point and the surface of its corresponding point, or the orientation of the normal. Higher weight was assigned to the x axis since it is the axis of the driving direction, producing the highest deviation between the point cloud pairs. The maximum correspondence distance was set to 20 cm.

Initially, the position of the first point cloud was geo-referenced (first point cloud of every path pass) using the total station and its pose was established using the roll, pitch and yaw values of the IMU sensor. Therefore, the frame of the first *target* point cloud (or model shape*)* was used as the main reference. Then, since the incoming *source* point cloud (or data shape) was mostly composed of soil points, the point cloud rigid registration was dominated by them. For that, an iterative reduction of soil points, from the *source* point cloud, was performed until a threshold value was reached. The threshold value was calculated using the Eq. (6):

$$number\ of\ points_{threshold} = \frac{number\ of\ points_{source}}{2 \times plant\ points\ \%} \tag{6}$$

where the number of points of the *threshold* point cloud is a reduction of the number of points of the incoming *source* point cloud, which is inversely proportional to twice the percentage of the plant points. The threshold point cloud contain plant and soil points, but some soil point were removed from the bottom to the top by the RANSAC to minimize the error in ICP registration. Afterwards, a filtering was performed in order to remove the invalid points or NaN (Not a Number) values.

The purpose of this equation was to balance the plant/soil ratio equal to one, so that the filtered *source* point cloud (from now on: *threshold* point cloud) contained approximately the same number of plant and soil points; it was found that the ICP registration was more accurate with the plat/soil ratio equal to one. In the case of the dataset used in this research, the approximate number of points of the incoming *source* point cloud was equal to ~800,000 points, where ~5% belonged to plant points and the remaining ~95% belonged to soil points. By substituting in the Eq. (6), the resulting number of points of the *threshold* point cloud was equal to ~80,000 points. Once the RANSAC
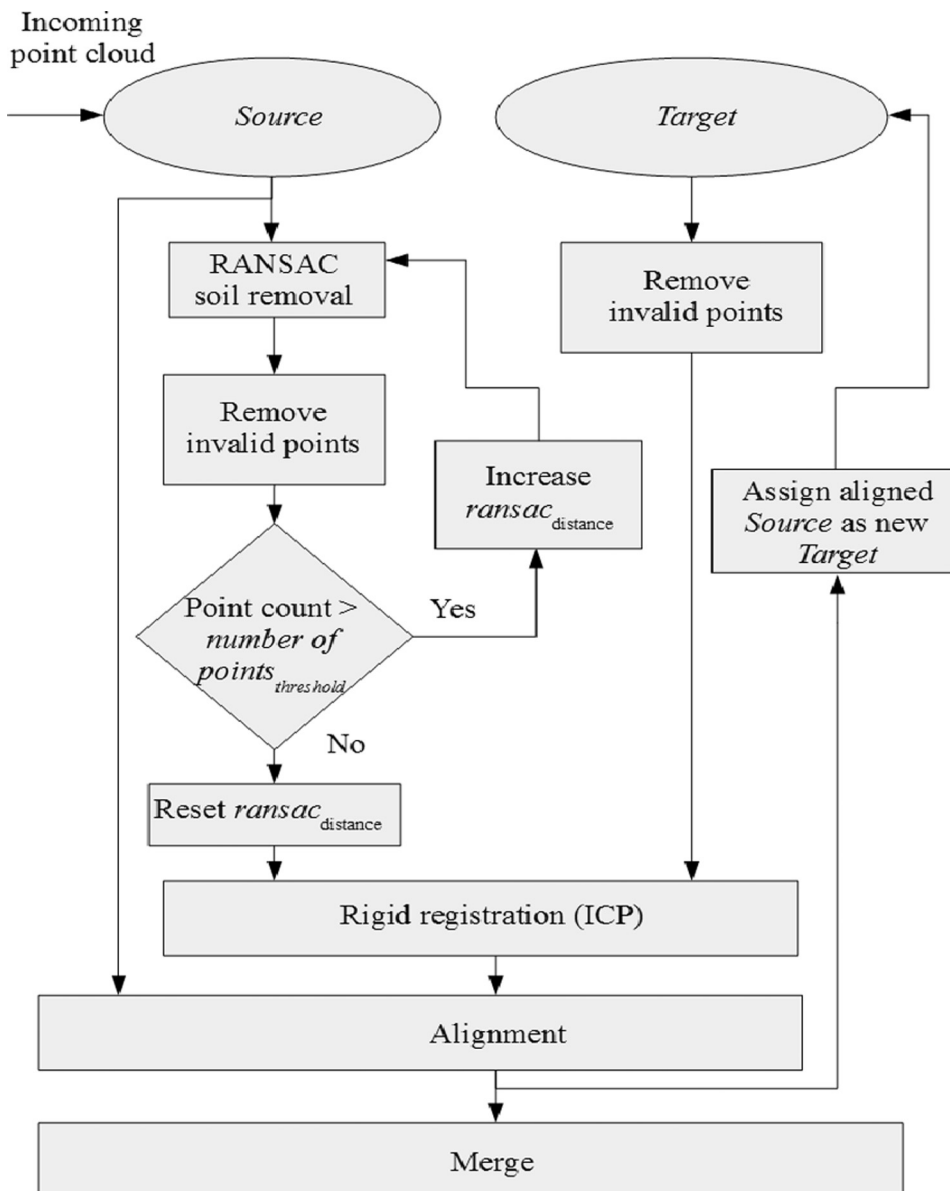

Fig. 4. Complete workflow of the adaptive point cloud rigid registration and stitching process.

plane fit algorithm was applied to obtain the plane model, the maximum distance $ransac_{distance}$ from the inliers (soil points) to the plane model was iteratively increased by a step of 5 mm, and therefore more soil points were removed. When the threshold value was reached, the $ransac_{distance}$ was reset to zero and the transformation that registered the filtered *threshold* point cloud to the *target* point cloud was estimated in the rigid registration process.

In order to accelerate and reinforce the transformation estimation, a *k*-d tree was used for the search of the nearest neighbours. Additionally, the point cloud pair (*target* and *source*) was weighted considering other criteria that described the similarity of the correspondence like the normals of the points in the moving direction (*x* axis) of the robotic platform. Then, using the estimated rigid transformation, the *source* point cloud was aligned to the reference defined by the first *target* point cloud. After that, the point cloud pair was merged. Finally, the *source* point cloud was assigned as the new *target*, and the next incoming point cloud was assigned as the new *source*. This process was performed iteratively until all the point clouds were stitched together.

2.3.4. Point cloud segmentation methodology

After the stitched point cloud were obtained, and for further analysis, the following processes was applied for plant/soil segmentation:

1. Voxel grid filter: The grid average down sample method was used to reduce the high density of points of the stitched point cloud. A grid step of 3 mm was used as, this value determined the size of the voxel where the inliers were merged into one single point.
2. Nearest neighbour filter: A nearest neighbour filter was applied by taking into consideration the standard deviation from the mean of the average distance to neighbours of all points, which was set equal to one; and the number of nearest neighbours, which was set equal to 10.
3. RANSAC: A plane fit algorithm was applied to each crop row to obtain the plane model of the soil points, by setting the $ransac_{distance}$ equal to 70 mm for plant/soil segmentation.

2.3.5. Point cloud validation methodology

The validation methodology for the accuracy of the maize plant reconstruction and alignment relative to the ground truth follows the next steps:

1. Project a parallel plane with a normal distance of 50 mm above the RANSAC plane model.
2. Segment the plant point cloud by keeping the points lying between the RANSAC and the parallel plane and removing the ones lying above the parallel plane.
3. Obtain the *x* and *y* values of all the inliers and perform an heuristic *k*-means clustering. Since some of the plants of the crop rows died while some others where very thin or small to be detected by the 3-D imaging acquisition system, the number of clusters in the data was reduced to a value visually selected based on the lateral view of the reconstructed crop rows.
4. Manually compare the clusters with the ground truth. A polynomial curve fitting of first order was performed considering the least absolute residuals (LAR), to minimize the effect of extreme values on the fit, in order to visualize the alignment of the clusters compared with the ground truth.

3. Results and discussion

After the point cloud paths were stitched in ROS, they were saved in ".bag" files. Then, the Robotics System Toolbox™ of MATLAB R2016b was used as an interface for exporting the stitched point clouds from ROS to MATLAB for further analysis. The toolbox provides programming classes and functions to export and filter ".bag" files.

3.1. Point cloud segmentation

The result of the registration and stitching of all point clouds acquired along path 2 is shown in Fig. 5a while going and Fig. 5b while returning. It can be seen that due to the complex plant architecture, the
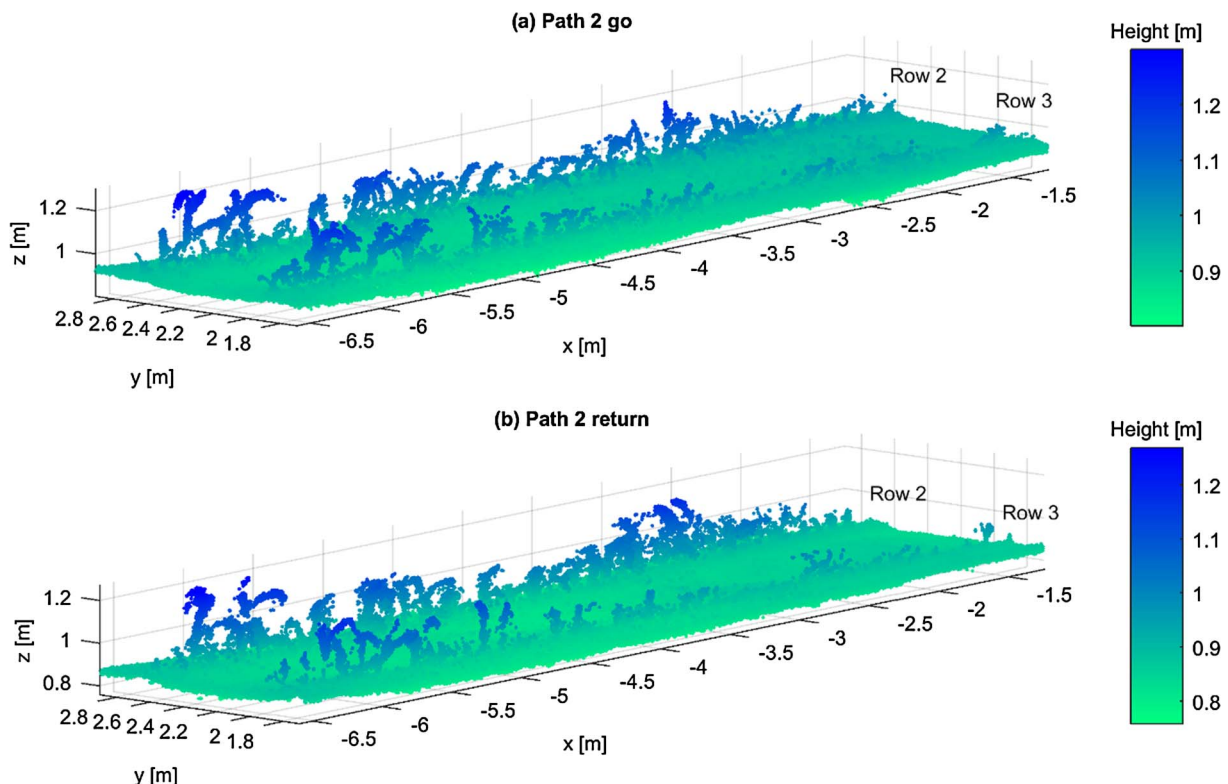

Fig. 5. Complete path 2 registration and stitching where the crop rows 2 and 3 are clearly recognisable (a) go; and (b) return.

Table 3
Point cloud reduction from path point clouds for plant point extraction.

Direction	Path	Raw data	Voxel grid filter	Noise filter	Plant points
Go	1	12,516,043 (100%)	3,472,891 (27.7%)	3,368,255 (26.9%)	157,707 (1.2%)
	2	9,150,946 (100%)	3,027,440 (33.0%)	2,908,181 (31.7%)	69,456 (0.7%)
	3	9,243,466 (100%)	3,120,200 (33.7%)	2,975,533 (32.1%)	56,777 (0.6%)
	4	9,066,749 (100%)	3,221,071 (35.5%)	3,119,758 (34.4%)	47,298 (0.5%)
Return	1	10,023,226 (100%)	3,135,542 (31.2%)	3,005,585 (29.9%)	108,465 (1.0%)
	2	10,568,564 (100%)	3,193,560 (30.2%)	3,076,468 (29.1%)	77,836 (0.7%)
	3	10,764,755 (100%)	3,386,178 (31.4%)	3,227,852 (29.9%)	47,490 (0.4%)
	4	8,601,601 (100%)	2,795,031 (32.4%)	2,665,274 (30.9%)	43,485 (0.5%)
Average		9,991,919	3,168,989	3,043,363	76,064

Table 4
Point cloud reduction of crop row point clouds for plant point extraction.

Direction	Crop row	Raw data	Voxel grid filter	Noise filter	Plant points
Go	1	6,524,620 (100%)	1,803,445 (27.6%)	1,744,851 (26.7%)	71,300 (1.0%)
	2	4,237,781 (100%)	1,425,371 (33.6%)	1,376,506 (32.4%)	50,029 (1.1%)
	3	4,913,167 (100%)	1,602,300 (32.6%)	1,517,289 (30.8%)	14,594 (0.2%)
	4	4,438,620 (100%)	1,577,788 (35.5%)	1,530,654 (34.4%)	28,649 (0.6%)
	5	4,628,129 (100%)	1,643,137 (35.5%)	1,586,639 (34.2%)	20,314 (0.4%)
Return	1	5,237,704 (100%)	1,599,316 (30.5%)	1,539,911 (29.4%)	54,540 (1.0%)
	2	5,075,875 (100%)	1,563,370 (30.8%)	1,514,102 (29.8%)	56,154 (1.1%)
	3	5,492,692 (100%)	1,630,232 (29.6%)	1,553,520 (28.2%)	18,363 (0.3%)
	4	4,336,064 (100%)	1,401,624 (32.3%)	1,353,474 (31.2%)	27,613 (0.6%)
	5	4,265,537 (100%)	1,393,260 (32.6%)	1,303,087 (30.5%)	17,103 (0.4%)
Total		49,150,189 (100%)	15,639,843 (31.8%)	15,020,033 (30.5)	358,659 (0.7%)

3-D shape reconstruction is slightly different when driving in the opposite direction; but still the maize leaves are clearly defined.

After the registration and stitching process, the resulting 3-D point cloud of every path is presented in Table 3 for both directions (go and return). The raw data column represents the reconstructed path containing two crop rows as seen in Fig. 3. After applying a voxel grid filter, a noise filter (near neighbours) and plant/soil segmentation by using the RANSAC fit algorithm, it can be seen that the points that belong to plants are just between 0.4% and 1.2% of the total amount of points. These results justify not just our statement regarding the dominance of soil points, but also the high density of points that needs to be handled with this TOF camera. According to Table 3, the average point cloud count of a scene, such as the ones depicted in Fig. 5, is roughly 10 million points where just 76,064 belonged to plants.

By dividing the path into two halves, in order to obtain the individual crop rows, the results presented in Table 4 are comparable with the ones of Table 3. The voxel grid filter eliminated ~70% of the points and the final plant points were also ~1%. Comparing these results with the paper written by Garrido et al. (2015) (they used 3 LIDARs) where the data provided by the vertical and inclined (45°) LIDARs were more than 90% of the points, and the rest where from the horizontally mounted LIDAR. The inclined LIDAR is the one that could be somehow comparable with the TOF camera, since they were both mounted very close together with the same angle as seen in Figs. 1 and 2. The average raw data from the inclined LIDAR was 60,720 points, in comparison, from the TOF camera the average was 4,915,018 points.

In the paper by Garrido et al. (2015), the vertical LIDAR was mounted in an advantageous position (see again Figs. 1 and 2), thus it was the one providing most of the useful data. If a LIDAR fusion is considered, the total number of raw data points coming from the inclined and vertical LIDARs together for every row (go and return) was 2,062,510 points. This value comes short compared to the 49,150,189 points coming from the TOF camera (see Table 4), which turns to be roughly 23 times larger.

Since the objective of this paper is to do a qualitative analysis of the plant 3-D shape reconstruction, it was necessary to compare the visual aspect of the generated point clouds from different perspectives. Fig. 6

shows a section of the crop rows 2 and 3, while the robotic platform was following path 2 going and returning. In the resulting point clouds, it can be seen that at a particular height, the occlusion starts to be evident as a void in the soil (Fig. 6c and e). Also, it can be noticed that the visual quality of the point cloud depends on the perspective relative to the driving direction. The leaves in Fig. 6d (e.g. plant 11) and Fig. 7c (e.g. plant 3) seem more defined compared with the ones in Fig. 6f and c, respectively. In other words, when the 3-D perspective view of the reconstructed plants is close to the one of the 3-D imaging acquisition system, the leaves appear more defined since they were facing the camera. Nevertheless, when the perspective view deviates from the one of the acquisition system, the leaves do not appear well defined with some voids, flying (veil) points or leaves with thick appearance clearly visible. Nevertheless, it is still possible to visually recognise the plants.

It is also interesting to compare the point clouds while going, returning and the fusion of both of them. In Fig. 6e and f, the point clouds of plant 1 (16 cm), 2 (24 cm), 9 (21 cm) and 10 (23 cm) are directly after (in the return direction) the tallest plants of their respective rows: plant 3 (47 cm) and plant 11 (35 cm). Therefore, they are barely noticeable compared with the point clouds of Fig. 6c and d, where plant 1, 2, 9 and 10 are more clearly defined. It can be inferred that plant occlusion is still an issue in some cases due to plant height heterogeneity and the driving direction of the 3-D imaging acquisition system.

The advantage of the TOF light volume technique is that some information can be obtained from small and thin objects. In Fig. 7b, plant 15 (height = 9 cm, stem width = 0.13 cm) can barely be seen, however, some 3-D information is obtained as seen in the magnification. Nevertheless, this point cloud is so small that can be lost in a filtering or segmentation process, like in Fig. 7f and h, were no green points are visible in the place where plant 15 actually is. This information provides an idea of the limits of this particular TOF camera.

Table 5 shows the number of plant points of every crop row going and returning and their fusion. A voxel grid filter was applied to maintain the same point density after fusion. Since the plant points are very close together, approximately 30% of the points were removed after the filter. Fig. 8b depicts the reconstruction of the whole field (Fig. 8a) going and returning with 246,182 plant and 23,832,277 soil points.
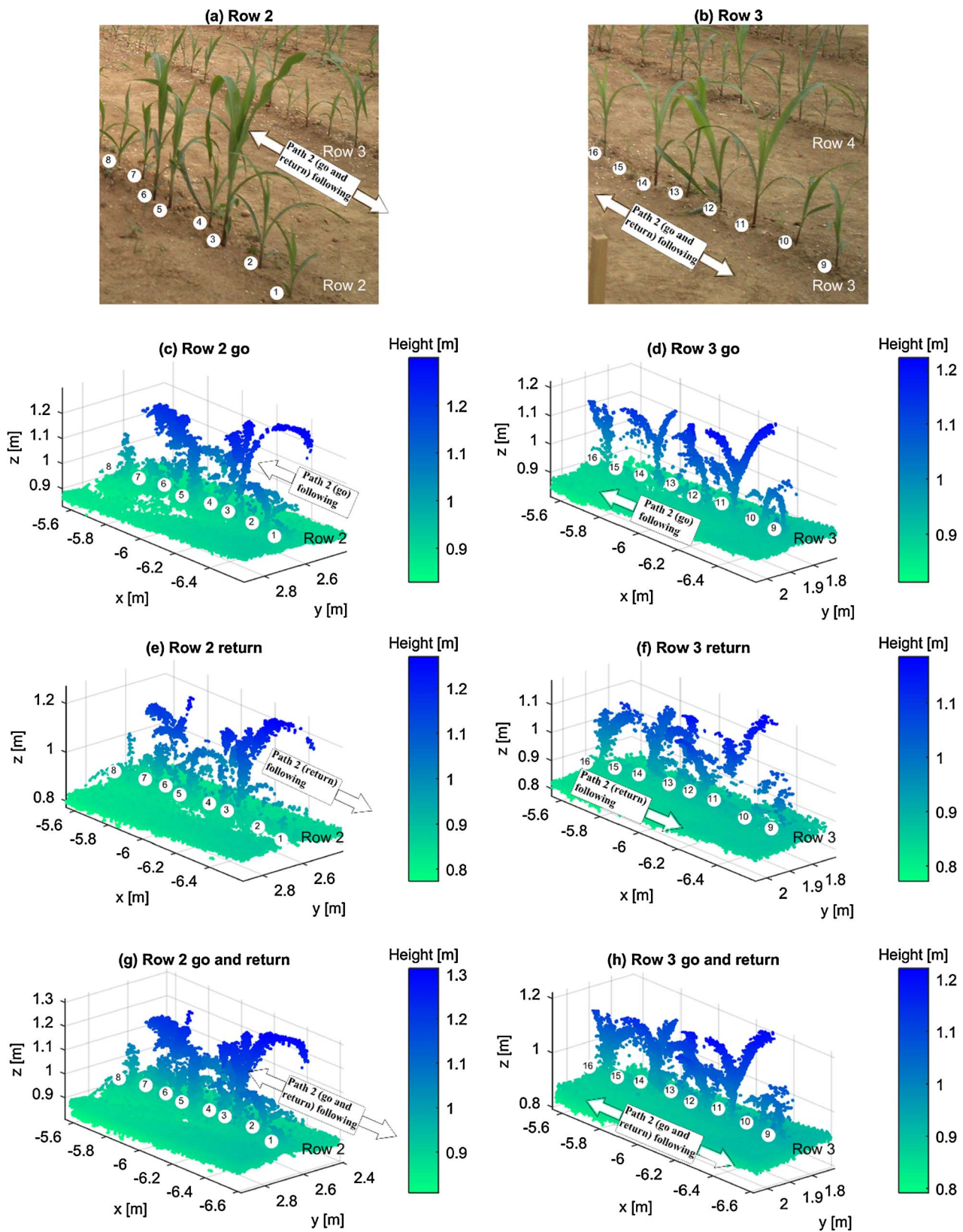
Fig. 6. Path 2 section. (a) row 2 RGB image; (b) row 3 RGB image; (c) row 2 go; (d) row 3 go; (e) row 3 return; (f) row 3 return; (g) row 2 fusion; (h) row 3 fusion.

3.2. Point cloud validation

The process of obtaining the ground truth was slow and tedious but was very useful to validate the maize plant reconstruction with high precision. However, for practical applications, a better option would be to geo-reference artificial markers such as metal tubes that are pressed or hammered into the soil. These markers need to be placed particularly at the beginning and the end of every crop row, and in between them in the case of long crop rows. The markers would be also reconstructed and they would provide information not just to validate the
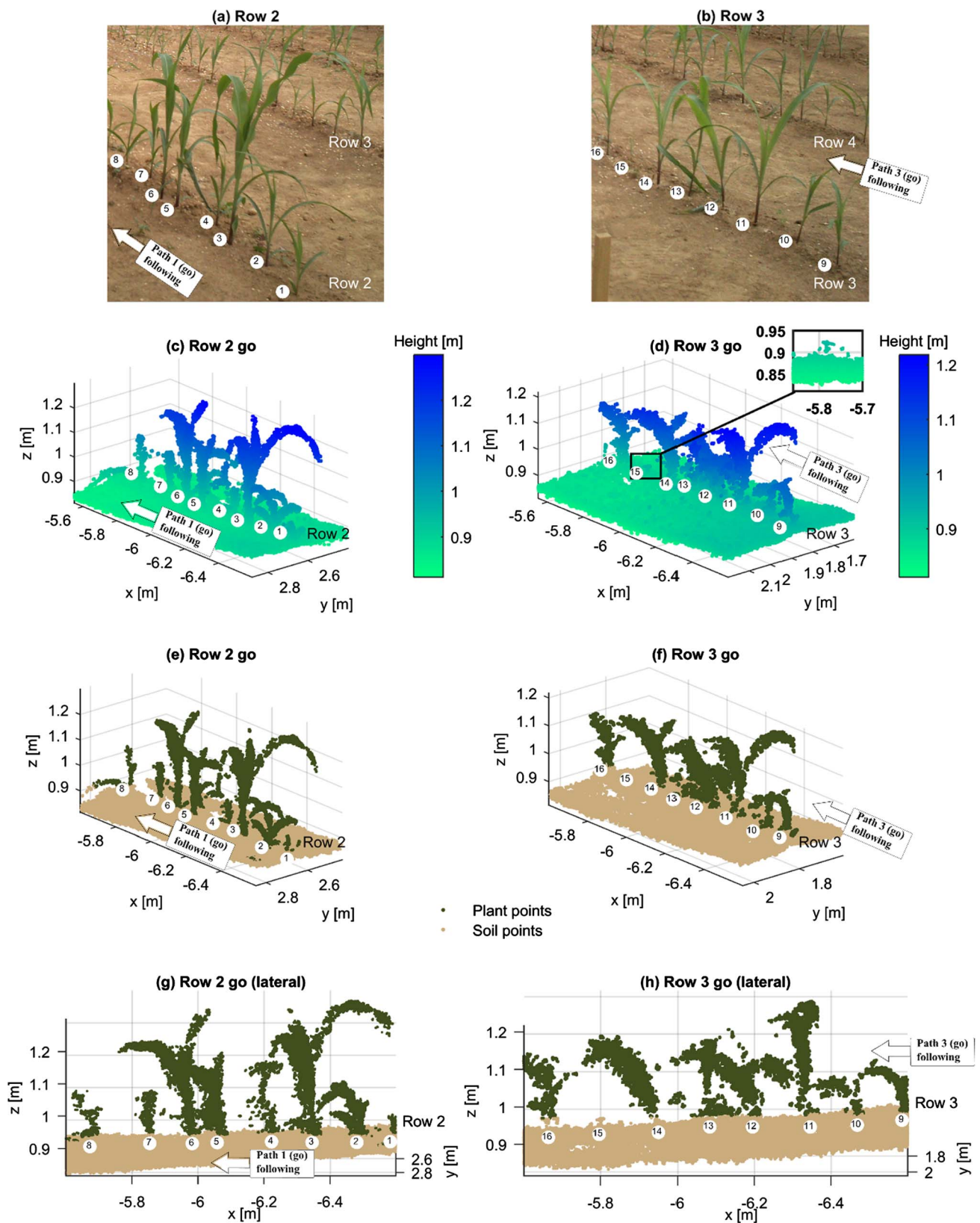
Fig. 7. Path 1 section. (a) row 2 RGB image; (b) row 3 RGB image; (c) row 2 go; (d) row 3 go; (e) row 3 return; (f) row 3 return (g) row 2 return lateral (h) row 3 return lateral.

reconstruction, but also to recalibrate the registration and stitching algorithm in the case that it starts accumulating errors.

In this research, in order to validate the accuracy of the 3-D plant shape reconstruction, an evaluation of the *x* and *y* position of the stems from the reconstructed crop rows, relative to the ground truth,

was performed. For that, a segmentation of the plant points (see Fig. 9a) was done. The plant points above a 5 mm plane parallel to the RANSAC plane model were removed. The remaining plant points are shown in Fig. 9b. Then, a 2-D *k*-means clustering was applied to those points. Since some plants were not detected because of reasons such

Table 5
Plant point cloud fusion and down sample.

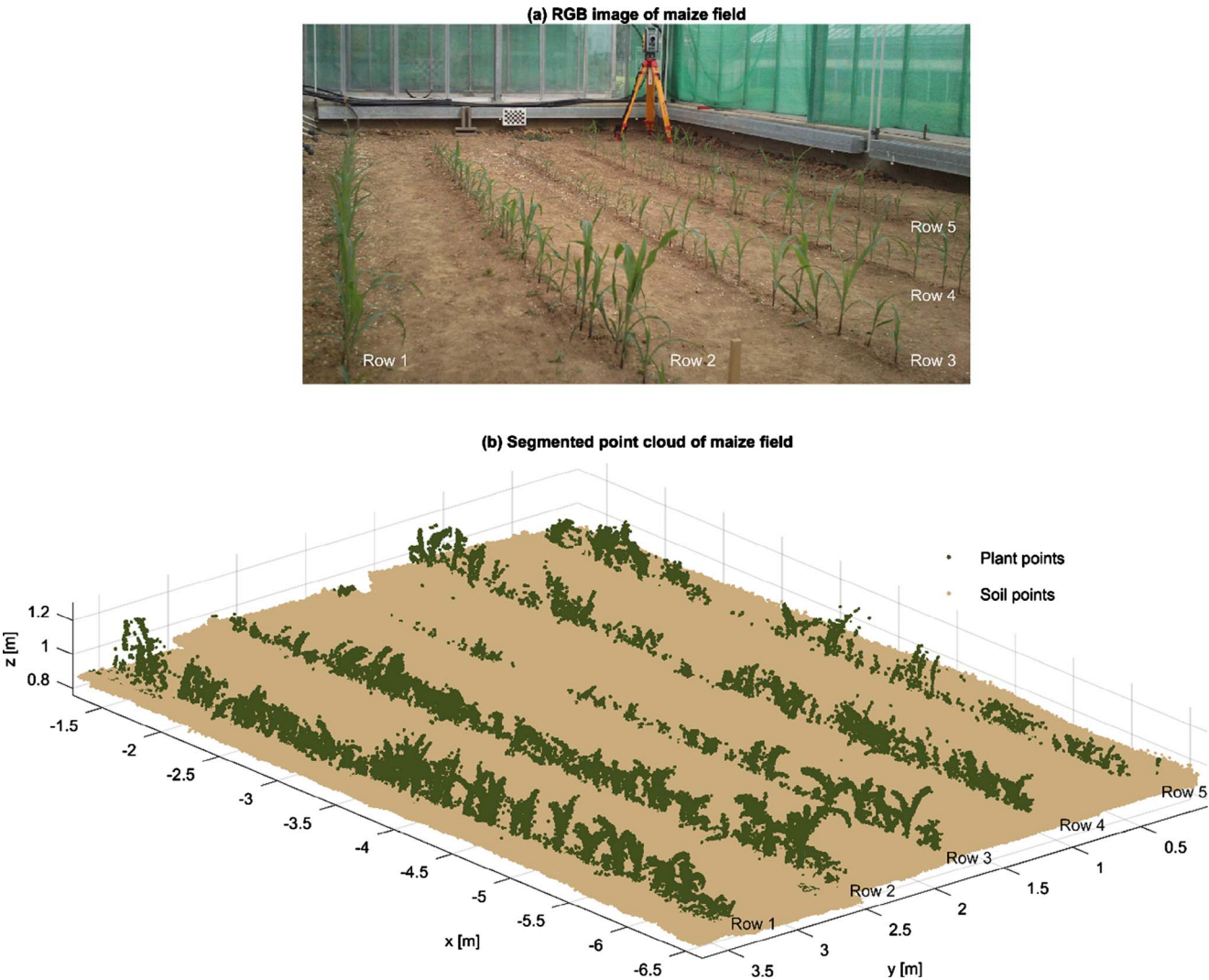| Row	Go	Return	Go + return "voxel grid filter"	Go + return	Root mean square error [m]
1	71,300 (56.6%)	54,540 (43.3%)	89,445 (71%)	125,840 (100%)	0.020
2	50,029 (47.1%)	56,154 (52.8%)	73,363 (69%)	106,183 (100%)	0.020
3	14,594 (44.2%)	18,363 (55.7%)	24,373 (73%)	32,957 (100%)	0.032
4	28,649 (50.9%)	27,613 (49%)	37,274 (66.2%)	56,262 (100%)	0.037
5	20,314 (54.2%)	17,103 (45.7%)	21,727 58%	37,417 (100%)	0.038
Total	184,886 (51.5%)	173,773 (48.4%)	246,182 (68.6%)	358,659 (100%)	

Fig. 8. Maize field depicted as: (a) RGB image and (b) down sampled and segmented 3-D reconstruction going and returning.

as failed germination, death, occlusion or sensor data acquisition limitations; the number of clusters was set to different values (depending on the row), while keeping 13 cm as the distance parameter between clusters. The number of clusters is related with the number of reconstructed plants, which is more apparent from the lateral perspective (Fig. 7g and h). Finally, the 2-D data of the clusters was used to perform a least absolute residuals (LAR) fitting, to minimize the effect of extreme values on the fit, with a polynomial curve of first order. The resulting fit was used as a reference of the alignment in the *y* axis relative to the ground truth.

The alignment in the *y* axis is evident in Fig. 10, but it is not in the *x* axis, for that, particular attention must be paid to the first and last plants (left and right-side end of the figure, respectively) of each row. The first and last plants of row 2 where clustered with high precision

relative to the ground truth, but in row 3, just the first plant was clustered while the last was not. The failed clustering can be explained by the size of the plant, which was small and thin (height = 10 cm, stem width = 0.18 cm), additionally the height of the previous plant along the driving direction was 25 cm high, therefore some occlusion was affecting the 3-D data acquisition.

In order to evaluate the accuracy of the clustering in Fig. 10, the position in the *x* and *y* coordinates was compared with the ground truth. The mean and the standard deviation are shown in Table 6. It can be seen that the mean was in average 3.4 cm while the standard deviation was below ± 1.7 cm. The results in terms of detected stems, as a percentage of the total number of seedlings, closely agrees with the ones obtained by Reiser et al. (2016) which were roughly 60% in average.
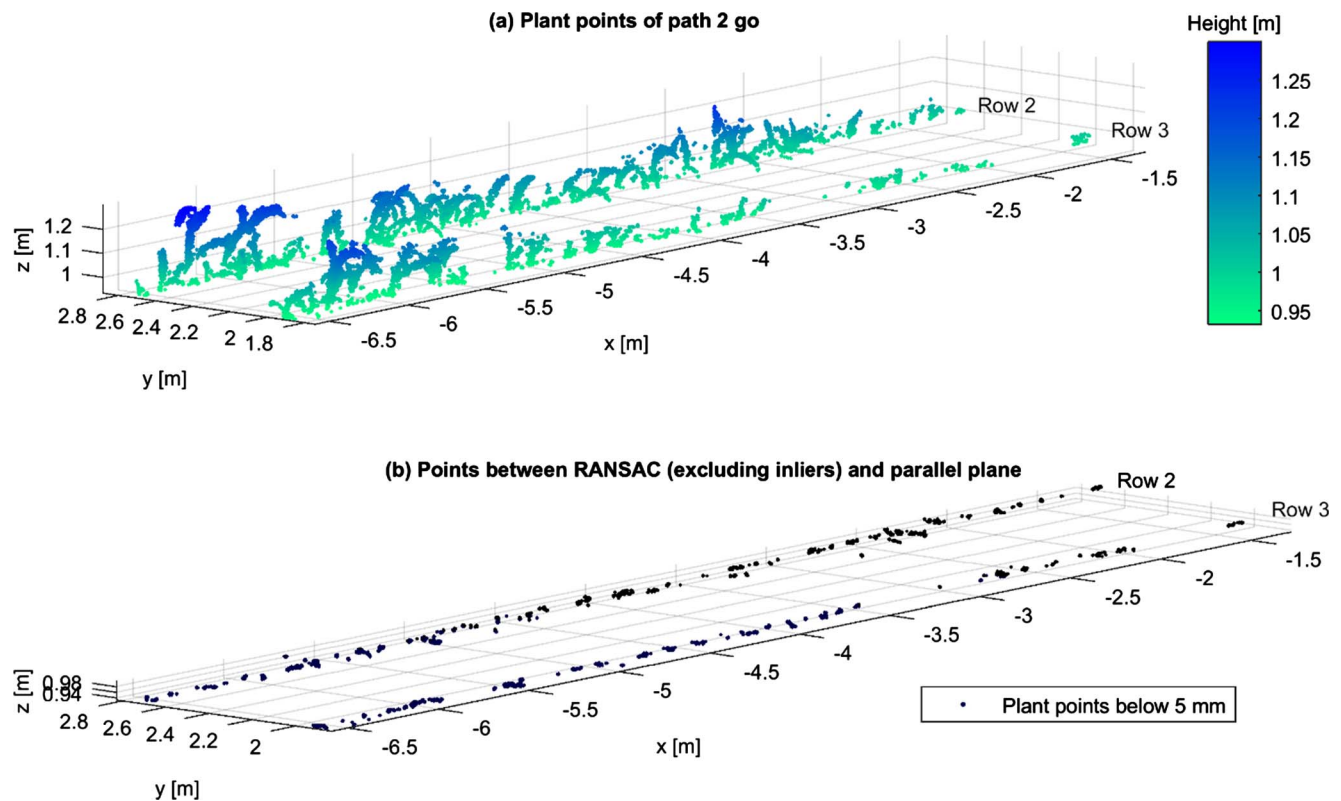
Fig. 9. Point cloud segmentation; (a) showing the plants after applying the RANSAC soil removal; (b) plant point cloud between RANSAC and parallel plane.
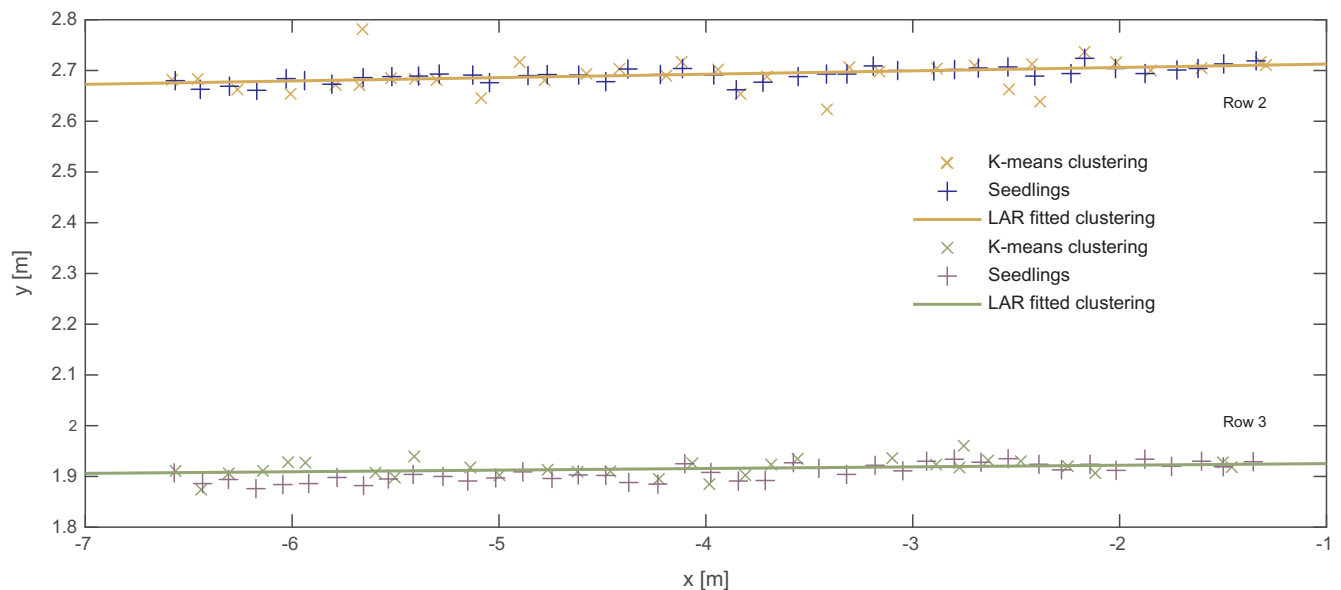

Fig. 10. Validation methodology showing the *k-means* clustering result (x) with a LAR fitting curve of the clustering and the seedling locations (+).

Jin and Tang (2009) also estimated the maize centre position, using a stereo vision system pointing perpendicularly downwards over the maize, reporting with their methodology a 74.6% detection of maize centres within 5 cm and 62.3% within 1 cm. However, they did not perform maize plant reconstruction, using instead the depth images from the stereo vision system. The time cost of the validation process, starting with an assembled path such as the one in Fig. 5, until the LAR fitting curve of the clustering was 28.12 s. The specific time cost of the *k-means* clustering and the LAR fitting was 18.45 and 5.47 s, respectively. The impact of this methodology is the demonstration that a georeferenced point cloud assembly is possible using a cost-effective TOF

camera. Although a high-resolution 3-D image was used in this research, using the resolution of the sensor image in combination with adequate filtering could be enough to perform the registration and stitching in real-time.

4. Conclusions

Within this research paper, we have described an adaptive methodology using a TOF camera for point cloud rigid registration and stitching. The resulting maize 3-D point clouds were highly dense and generated in a cost-effective manner.

Table 6
This table shows the result of the clustering compared with the seeding locations.

Direction	Crop row	Clusters [no.]	Correctly detected stems [no.]	False positives [no.]	Plants detected [%]	Mean [m]	Standard deviation [m]
Go	1	37	26	11	63	0.038	0.014
	2	35	27	8	66	0.022	0.008
	3	30	25	5	61	0.027	0.011
	4	35	24	11	59	0.044	0.017
	5	30	23	7	56	0.039	0.014
Return	1	37	24	13	59	0.042	0.016
	2	35	24	11	59	0.041	0.015
	3	30	13	17	32	0.045	0.013
	4	35	27	8	66	0.021	0.009
	5	30	26	4	63	0.024	0.013
Average		33.4	23.9	9.5	58.4	0.034	0.013

The validation of the methodology showed that the plants were reconstructed with high accuracies and the qualitative analysis showed the visual variability of the plants depending on the 3-D perspective view. However, independently of the view, the leaves were defined with qualities not seen until now with this type of sensor in maize plants. The results of this paper were also compared with the ones obtained with two LIDARs, where the point density was 23 times higher using the TOF camera. This methodology can be replicated in outdoor conditions since there have been already publications, cited within this research, relying on the same sensor in open field environments.

Further research directions should go to the individual plant level in order to evaluate the accuracy of the single plant 3-D reconstruction, as well as phenotyping properties such as plant and stem height, leaf area index (LAI), leaf angle, number of leaves or biomass.

Acknowledgments

Conflicts 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 16, 1–11. http://dx.doi.org/10.3390/s16070972.

Bentley, J.L., 1975. Multidimensional binary search trees used for associative searching. Commun. ACM 18, 509–517.

Besl, P., McKay, N., 1992. A method for registration of 3-D shapes. IEEE Trans. Pattern Anal. Mach. Intell. http://dx.doi.org/10.1109/34.121791.

Busemeyer, L., Mentrup, D., Möller, K., Wunder, E., Alheit, K., Hahn, V., Maurer, H., Reif, J., Würschum, T., Müller, J., Rahe, F., Ruckelshausen, A., 2013. BreedVision — a multi-sensor platform for non-destructive field-based phenotyping in plant breeding. Sensors 13, 2830–2847. http://dx.doi.org/10.3390/s130302830.

Butkiewicz, T., 2014. Low-cost Coastal Mapping using Kinect v2 Time-of-Flight Cameras.

Dhondt, S., Wuyts, N., Inzé, D., 2013. Cell to whole-plant phenotyping: the best is yet to come. Trends Plant Sci. 18, 428–439. http://dx.doi.org/10.1016/j.tplants.2013.04.008.

Dong, J., Burnham, J.G., Boots, B., Rains, G., Dellaert, F., 2017. 4D Crop Monitoring: Spatio-Temporal Reconstruction for Agriculture. In: IEEE International Conference on Robotics and Automation, Singapore.

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. http://dx.doi.org/10.1145/358669.358692.

Garrido, M., Paraforos, D.S., Reiser, D., Vázquez-Arellano, M., Griepentrog, H.W., Valero, C., 2015. 3D maize plant reconstruction based on georeferenced overlapping LiDAR

point clouds. Remote Sens. 7, 17077–17096. http://dx.doi.org/10.3390/rs71215870.

Hoffmeister, D., Waldhoff, G., Curdt, C., Tilly, N., Bendig, J., Bareth, G., 2013. Spatial variability detection of crop height in a single field by terrestrial laser scanning. In: Stafford, J.V. (Ed.), Precision Agriculture '13. Wageningen Academic Publishers, Wageningen, pp. 267–274. http://dx.doi.org/10.3920/978-90-8686-778-3_31.

Jin, J., Tang, L., 2009. Corn plant sensing using real-time stereo vision. J. F. Robot. 26, 591–608. http://dx.doi.org/10.1002/rob.

Kataoka, T., Kaneko, T., Okamoto, H., Hata, S., 2003. Crop growth estimation system using machine vision. In: IEEE/ASME International Conference on Advanced Intelligent Mechatronics. IEEE, Kobe, Japan, pp. 1079–1083. doi: http://dx.doi.org/10.1109/AIM.2003.1225492.

Kazmi, W., Foix, S., Alenyà, G., Andersen, H.J., 2014. Indoor and outdoor depth imaging of leaves with time-of-flight and stereo vision sensors: Analysis and comparison. ISPRS J. Photogramm. Remote Sens. 88, 128–146. http://dx.doi.org/10.1016/j.isprsjprs.2013.11.012.

Klose, R., Penlington, J., Ruckelshausen, A., 2009. Usability study of 3D time-of-flight cameras for automatic plant phenotyping. Image Anal. Agric. Prod. Process. 69, 93–105.

Lachat, E., Macher, H., Landes, T., Grussenmeyer, P., 2015a. Assessment and calibration of a RGB-D camera (Kinect v2 Sensor) towards a potential use for close-range 3D modeling. Remote Sens. 7, 13070–13097. http://dx.doi.org/10.3390/rs71013070.

Lachat, E., Macher, H., Mittet, M., Landes, T., Grussenmeyer, P., 2015. First experiences with Kinect v2 sensor for close range 3D modelling. In: The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. ISPRS, Avila, Spain, pp. 93–100. doi: http://dx.doi.org/10.5194/isprsarchives-XL-5-W4-93-2015.

Mai, C., Zheng, L., Li, M., 2015. Rapid 3D reconstruction of fruit tree based on point cloud registration. Trans. Chinese Soc. Agric. Eng. 31, 137–144.

Microsoft, 2017. Kinect hardware [WWW Document]. Kinect Hardw. key Featur. benefits. URL https://developer.microsoft.com/en-us/windows/kinect/hardware (Accessed 7.19.17).

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., Martín-López, J.M., Memic, E., Vázquez-Arellano, M., Brandner, S., Griepentrog, H.W., 2017. 3D imaging with a sonar sensor and an automated 3-axes frame for selective spraying in controlled conditions. J. Imaging 3. http://dx.doi.org/10.3390/jimaging3010009.

Reiser, D., Vázquez-Arellano, M., Garrido Izard, M., Griepentrog, H.W., Paraforos, D.S., 2016. Using assembled 2D LiDAR data for single plant detection. In: MCG 2016–5th International Conference on Machine Control & Guidance. Vichy, France.

Ritchie, S., Hanway, J., Benson, G., 1992. How a Corn Plant Develops. Ames, USA.

Rosell-Polo, J.R., Gregorio, E., Gene, J., Llorens, J., Torrent, X., Arno, J., Escola, A., 2017. Kinect v2 sensor-based mobile terrestrial laser scanner for agricultural outdoor applications. IEEE/ASME Trans. Mechatronics 4435, 1. http://dx.doi.org/10.1109/TMECH.2017.2663436. 1–1.

Rusu, R.B., Cousins, S., 2011. 3D is here: point cloud library. IEEE Int. Conf. Robot. Autom. 1–4. http://dx.doi.org/10.1109/ICRA.2011.5980567.

Vázquez-Arellano, M., Griepentrog, H.W., Reiser, D., Paraforos, D.S., 2016a. 3-D imaging systems for agricultural applications - a review. Sensors (Switzerland) 16. http://dx.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.), Intelligente Systeme - Stand Der Technik Und Neue Möglichkeiten. Gesellschaft für Informatik e.V. (GI), Osnabrück, Germany, pp. 213–216.

Weiss, U., Biber, P., 2011. Plant detection and mapping for agricultural robots using a 3D LIDAR sensor. Rob. Auton. Syst. 59, 265–273. http://dx.doi.org/10.1016/j.robot.2011.02.011.

Wiedemeyer, T., 2015. iai_kinect2. Inst. Artif. Intell.

Yin, X., Noguchi, N., 2013. Development of a target following system for a field robot. Comput. Electron. Agric. 98, 17–24.