

Plant detection and mapping for agricultural robots using a 3D LIDAR sensor

Ulrich Weiss*, Peter Biber

Corporate Sector Research and Advance Engineering, Robert Bosch GmbH, Robert Bosch Straße 19, 71701 Schwieberdingen, Germany

ARTICLE INFO

Article history:

Available online 11 March 2011

Keywords:

Individual plant detection
Plant mapping
3D LIDAR sensor
Agricultural robotics

ABSTRACT

In this article, we discuss the advantages of MEMS based 3D LIDAR sensors over traditional approaches like vision or stereo vision in the domain of agricultural robotics and compare these kinds of sensors with typical 3D sensors used on mobile robots. Further, we present an application for such sensors. This application deals with the detection and segmentation of plants and ground, which is one important prerequisite to perform localization, mapping and navigation for autonomous agricultural robots. We show the discrimination of ground and plants as well as the mapping of the plants. Experiments conducted using the FX6 LIDAR by Nippon Signal were carried out in the simulation environment Gazebo, with artificial maize plants in the laboratory and on a small maize field. Our results show that the tested plants can be reliably detected and segmented from ground, despite the use of the low resolution FX6 sensor. Further, the plants can be localized with high accuracy.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

Precision agriculture is about applying the right treatment, at the right place and at the right time [1]. To be really precise the treatment should be carried out adapted for each crop plant. But, what is the right treatment, when and where should it be carried out? For making such decisions it is necessary to collect, store and process crop field data at a subfield level. Nowadays it is a common procedure to perform this task by human labor using random samples.

The labor-intensive nature of precision agriculture practices brings a great need for automation [2]. Hence, one main objective of precision farming is the automation of the rating, storage, management and decision making processes by using autonomous vehicles. In future, such vehicles are needed for soil sampling, crop scouting, and real-time data collection [2]. Because of the huge size and heavy weight of conventional field machinery and the resulting compacting of the field, small robots are preferable. Furthermore, small robots are better suited for individual plant care, such as selective crop harvesting and individual crop rating as well as precise spraying, fertilization and weeding, to reduce costs, fertilizers and pesticides.

The work presented here is a part of the publicly funded project BoniRob [3], a cooperation of academic and industrial partners, with the goal to develop an autonomous mobile agricultural robot. The key objective of BoniRob is the autonomous repeating phenotyping of individual plants at different days, for that it is necessary to detect and map single plants. The automatic

phenotyping and mapping of all plants in a field will be a revolutionary change in the methods of field trials. Moreover, the availability of a robust crop scout will offer options for other field applications, such as stated in the previous paragraph.

A basic component of the crop scout is the navigation module, based on GPS localization and 3D environment sensing. The task of the navigation module is to guide the robot safely and autonomously over the field as well as to localize it even in unknown environments. For that purpose, in mobile robotics it is necessary to be sure about the position and the surroundings of the vehicle. In our case, we are detecting crop rows or even individual plants. We detect the individual plants because of different reasons: first for the navigation—the robot must follow a crop row without harming the plants, second for the localization—the rows and the plants can be used as natural landmarks and third for the mapping of the plants—we have to retrieve the individual plants because the phenotyping should be performed at different days to track the growth stages. Fig. 1 shows our test robot sensing the crop with a 3D laser sensor and a point cloud of an artificial crop row, also showing the detected plants.

At present, 2D vision based approaches are common for environment sensing in agricultural robotics. Some groups are using stereo vision to receive 3D data, but 3D laser sensors are not in the focus of the community. But, one obvious advantage of 3D laser sensors over 2D vision and stereo vision is the availability of reliable ranging data, which relieves the object detection and the object localization. Other advantages over 2D vision and other 3D sensor technologies are the robustness against illumination and atmospheric conditions, which enables the robot to operate reliable at any weather conditions even 24 h a day.

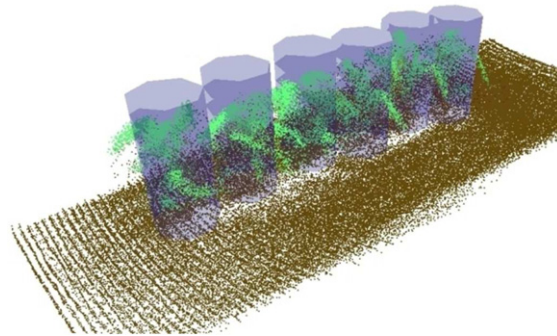
In the first part of this paper, we discuss the state-of-the-art in environment sensing in agricultural robotics and introduce the

* Corresponding author. Tel.: +49 711 811 24882; fax: +49 711 811 5183157.

E-mail addresses: ulrich.weiss@de.bosch.com (U. Weiss), peter.biber@de.bosch.com (P. Biber).



(a) Robot sensing maize plants.



(b) Point cloud and detected maize plants.

Fig. 1. In the picture (a) one can see our test robot on a maize field sensing the plants with the FX6 3D LIDAR sensor. The image (b) shows a point cloud of an artificial maize row, constructed using several laser scans. The points detected as ground are colored in brown; the points belonging to the plants are colored in green; the detected maize plants are marked using purple cylinders. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Micro-Electro-Mechanical Systems (MEMS) mirror based FX6 3D light detection and ranging (LIDAR) sensor by Nippon Signal [4]. We also compare this sensor with other state-of-the-art 3D sensor technologies suitable for mobile robotics. The second part of this paper is focused on an algorithm performing individual plant detection and mapping using a 3D LIDAR sensor on a mobile robot. Our proposed algorithm is based on an approach detecting the ground to segment the point cloud into soil and other objects. In the remaining point cloud it searches for clusters representing plants using a model of a crop row. According to the detected ground plane the plant localization and mapping is performed. In this paper we show that this task can be reliably carried out using the 3D-LIDAR sensor with its resolution of 59×29 rays. According to [5] the required localization accuracy for the task of individual plant treatment is about 1 cm. To achieve almost this precision we are applying sensor fusion for the global positioning of the robot using a real-time-kinematic (RTK)-GPS device, the odometer and an inertial measurement unit (IMU). The accuracy of this sensor fusion is about 2 cm, which has shown in our experiments to be sufficiently precise for our task.

This paper is organized as follows. In Section 2 we discuss the state-of-the-art of environment sensing in the domain of agricultural robotics and compare typical 3D sensors used on mobile robots with the FX6 sensor. The algorithm performing the ground and plant detection as well as the mapping is introduced in Section 3. Section 4 shows the experimental results. Finally, in Section 5, the conclusions and the future works are presented.

2. Environment sensing in agricultural robotics

2.1. State-of-the-art

If one is to believe [6], the most promising system of relative positioning is computer vision using cameras. Indeed, at present, for environment sensing in agricultural robotics and for automatic guidance in crop rows most approaches recently traced are 2D vision based. The main challenge here is to interpret the images to find a guidance directrix, e.g., the position and orientation of the crop rows relative to the vehicle [2].

There are many publications dealing with this topic, e.g. in [7] the authors present a vision based method for crop row recognition using the Hough Transform which is also robust against the presence of weed. By using a near-infrared filter on a gray-scale camera they get a high contrast image where living plant material and soil is easily discriminated. Tellaëche et al. [8] combine Support Vector Machines and k-means Fuzzy Logic for detecting weed in

cereal crop cultivations at early growing stages. They partition the image into cells, each described by two area-based attributes measuring the relation between crop and weed. John Billingsley introduced in [9] an overview over agricultural applications for 2D machine vision by several researchers, for task such as weed detection and eradication, automatic vehicle guidance, visual grading of produce, and identification and segregation of animal species.

Stereo vision systems are also known in agricultural applications, like [10] introduce two techniques for corn plant and population sensing, namely an active and a passive approach. In [11] stereo vision is used for the mapping of an agricultural environment like apple trees. Another application using stereo vision in this domain, proposed in [12], is obstacle detection to enhance safety by detecting persons in front of the vehicle.

One big disadvantage using vision or stereo vision is the influence by changing lighting conditions, which are not uncommon in outdoor applications like agriculture [2]. Laser rangefinder are a much more reliable technology in such situations, but up to now applications with laser sensors are rare in this domain. E.g., Satow et al. [13] developed an automatic crop row guidance system for tractor mounted implements, with a 2D line laser able to detect the height and the position of crop rows. Applications using 3D LIDAR sensors are not known in this domain.

The above mentioned publications deal with crop row detection to guide vehicles or with crop/weed/soil-discrimination but not with single plant detection and mapping. In [14] an approach for detecting individual plants, by generating an image using a light curtain which is moved along a crop row, was presented. Using a RTK-GPS on a non-autonomous mobile unit and sensor fusion they are able to generate a map of the crop rows with the position of individual plants. With such a sensor single plants can be detected, but the image generation depends on accurate position data, and because of the line characteristic of the sensor an image of a plant is not available until the vehicle pasts the object completely with the curtain. Another disadvantage is the necessity that the curtain has to be mounted at a low level to sense a whole plant, but at this level it may touch the plants. This sensor is also not able to look forward, which is required for the navigation task. In contrast, the method presented in this paper is able to detect individual plants up to a distance of 3 m in front of the robot, depending on the sensor mounting position.

2.2. 3D sensors in mobile robotics

Until today, most mobile robot implementations in indoor environments rely on 2D sensors for building maps, self-localization and collision avoidance. This is justified to some extent

because most applications were developed for ground robots, which inherently move in a 2D space [15]. However, these approaches will not work in outdoor applications like agricultural robotics where there are no flat grounds or upright walls. Here, 3D sensors promise to give the required information to perform mapping, self-localization, object detection, etc. The advantages of 3D sensors for object detection and mapping are obvious. On one hand it is much easier to build a 3D map out of 3D data and on the other hand measurement methods, which provide reliable distance information are useful for save and robust applications. Up-to-date 3D sensors like photon mixer devices (PMD) and LIDAR sensors also outputs intensity values, stereo vision cameras even provide color values.

Today three sensor technologies are mainly in use for 3D sensing on mobile robots: stereo vision, PMD time-of-flight cameras and laser sensors. Works like [16,15,17] compared and evaluated applications for 3D sensors on mobile robots, with the focus on indoor applications using stereo vision and PMD time-of-flight cameras. In contrast, we evaluated 3D laser sensors in outdoor environments, especially for agricultural robots.

To receive 3D data from normal cameras, typically stereo systems with two or more cameras or structure from motion techniques are used. Because of their passive operation mode, both have difficulties providing reliable data for navigation and mapping. To receive 3D data from stereo vision, structures in the images are required. Further, the precision and maximum depth is limited by the baseline between the cameras, and the quality of the distance values decreases very fast as depth increases. Big advantages of the stereo vision are, the high resolution images and the availability of color values. Further, stereo systems are relatively small and low priced, because one can use standard vision components, for instance embedded in almost every cell phone.

The semiconductor based PMD cameras are the latest technique and allow, additionally to the conventional intensity image to receive ranging information based on the time-of-flight by a modulated light source which is reflected by the environment and absorbed from this sensor [18]. The distances are calculated by the phase shift of the signal and the reflection intensity by the signal amplitude. By measuring the phase shift of the emitted infrared light the maximum distance is limited. Most light sources used by the time-of-flight cameras are buildup using LEDs with a modulation of 20 MHz which allows an ambiguous range measuring up to 7.5 m. Klose et al. [19] have analyzed the usability of such cameras for agricultural applications, like rating. They discovered that this sensor is capable for this application, but the currently available sensors still have some problems. The quality of the depth values depends on the color of the material reflecting the emitted light and some sensors have problems with moving objects by showing motion blur effects. General disadvantages of this kind of sensors are: the illumination with an active light source, which can be heterodyne by other light sources, particularly outdoors and varying depth values depending on the temperature of the sensor. But PMD cameras have not yet reached their full potential: from version to version the resolutions are increasing while the prices decrease and they are also getting more robust.

Among the 3D laser sensors the price range as well as the accuracy, resolution and frame rate is wide spread. A typical example for a 3D laser scanner on a mobile robot is the Kurt3D-project [20] which is equipped with a rotating 2D SICK laser. To receive consistent 3D data with such an approach a stop-and-go mode for traveling is necessary, because this 2D line sensors were not build for these applications. Some groups like Thrun et al. [21] use two 2D laser rangefinders for building 3D volumetric data. Other groups use highly precise and expensive 3D laser

sensors, like CYRAX [22], RIEGL [23] or Zoller+Fröhlich scanners [24]. These approaches share the properties of a heavy weight, high power consumption and expensive prices. The benefits are a high resolution and a large field of view.

2.3. The FX6 3D LIDAR sensor

In contrast to the highly precise and expensive laser sensors, the FX6 (Fig. 2(a)) sensor [4] is lightweight and has no rotational mechanical parts, but a low resolution. It measures the distance and reflectance intensity using an infrared pulsed laser light (Class 1) with up to 16 fps, determining the time-of-flight at a smallest unit of about 30 ps, which leads to a ranging precision of about 1 cm. The laser beam is reflected by a mirror oscillating independently into two directions. The combination of them results in a Lissajous figure shown in Fig. 2(b). The direction of the laser beam and the field of view of the photodiode can be concentrated on each measuring point respectively, thus it is possible to receive a signal even under strong scattering light conditions and keep the noise level minimized. This allows to measure the distance even in the sunlight. Fig. 3 is showing a picture of an artificial maize row and the corresponding depth and reflectance intensity image provided by the 3D-LIDAR as well as the point cloud constructed using the depth and intensity values.

Table 1 compares a typical stereo vision camera and a PMD sensor with the 3D-LIDAR sensor, used by us. The table shows, that besides the resolution, power consumption, weight, maximum range and frame rate the properties of the sensors are comparable. The main advantage of the LIDAR sensor, compared with both other technologies, is the reliability under changing lighting conditions, e.g., the sensor is not influenced by the sunlight and by other weather conditions, which was shown by our outdoor experiments. Using a LIDAR sensor it is even possible to operate 24 h a day and handle conditions like light fog and dust. In agricultural environments such conditions may occur.

In contrast to most stereo vision cameras, the LIDAR sensor calculates the ranging information on the sensor, so no time-consuming calculations has to be done on the computer. The precision of the distance values are also independent from the measured distances. In contrast to PMD cameras, the maximum distance is only limited by the intensity of the laser beam, because the sensor measures the time-of-flight and not the phase shift. Compared to line sensors like 2D laser scanners or light curtains a whole plant can be sensed at once using only one frame. In contrast to the light curtain it is possible to mount the 3D sensor looking forward, which relieves the navigation task.

According to Edan et al. [2] the most reliable technology is the laser rangefinder, but currently its use is limited to research vehicles platform due to the high costs. But by using the MEMS technology these sensors are potentially low cost, and in the near future such sensors will have higher resolutions and frame rates. They are even more mechanical robust than conventional laser sensors, because they do not consist of any rotational mechanical parts. Our experiments show the usability of these sensors for outdoor applications and that by using the FX6, even with its relatively low resolution, we are able to detect individual plants and their positions reliable.

3. Plant detection and mapping

In agricultural robotics many applications for 3D sensors are imaginable. Here, we focus on the task of plant detection and mapping to support the navigation of an autonomous agricultural robot. Our algorithm discriminates the ground and plants using a

Table 1
Characteristics of three typical 3D sensors.

| Sensortype/sensor name | Stereo vision/Bumblebee2 | PMD / SR-3000 | Laser / FX6 |
|------------------------|--|--|--|
| Manufacturer | Point gray | Mesa imaging | Nippon signal |
| Mode of operation | passive | active (illumination) | active (laser) |
| Range [m] | not specified | 7.5 | 16 |
| Precision [mm] | not specified | 22.5 (± 3 sigma) | 80 (± 3 sigma) |
| Resolution [pixel] | 640 \times 480 | 176 \times 144 | 29 \times 59 |
| Field of View [°] | 43 or 66 (HFOV) | 47.5 \times 39.6 (HFOV/VFOV) | 50 \times 60 (HFOV/VFOV) |
| Frame Rate [fps] | 48 | 25 | 16 |
| Output | x, y, z, color (RGB) | x, y, z, intensity | x, y, z, intensity |
| Size [mm] | 157 \times 36 \times 47.4 | 50 \times 67 \times 42.3 | 95 \times 64 \times 142 |
| Weight [kg] | 0.35 | 1.1 | 1 |
| Power consumption [W] | 2.5 (12V, ca. 0.2A) | 12–18 (12V, ca. 1A) | 7.2–5 (12V, ca. 0.5A) |
| Connection | FireWire | USB 2.0 | USB 2.0 |
| Approx. Price [EUR] | 1,800 | 4,000 | 5,000 |
| References | www.ptgrey.com | www.mesa-imaging.ch | www.signal.co.jp |

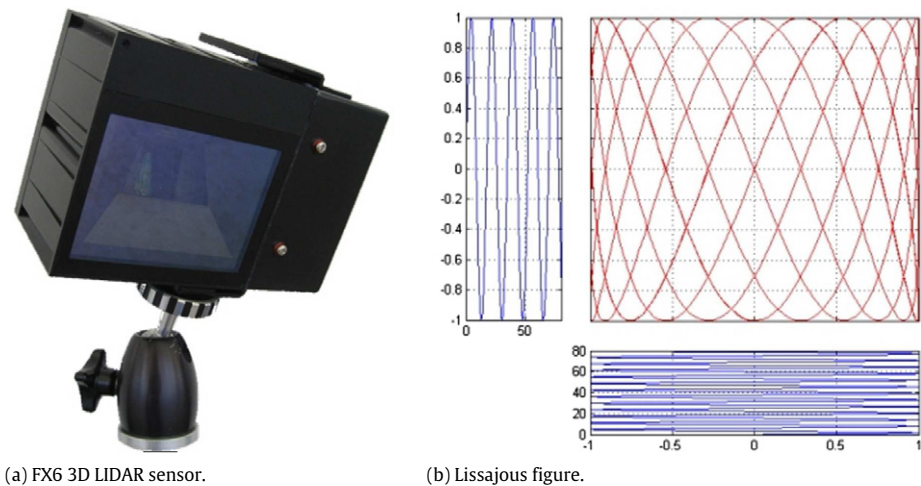


Fig. 2. The picture (a) shows the FX6 sensor mounted on a tripod. On image (b) an example of a Lissajous figure (red curve), described by the laser beam reflected from the MEMS mirror, is shown. This figure results from the combination of two heterodyne sine oscillations (blue curves). Using this figure for each frame the whole image with 59 \times 29 pixels is covered. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

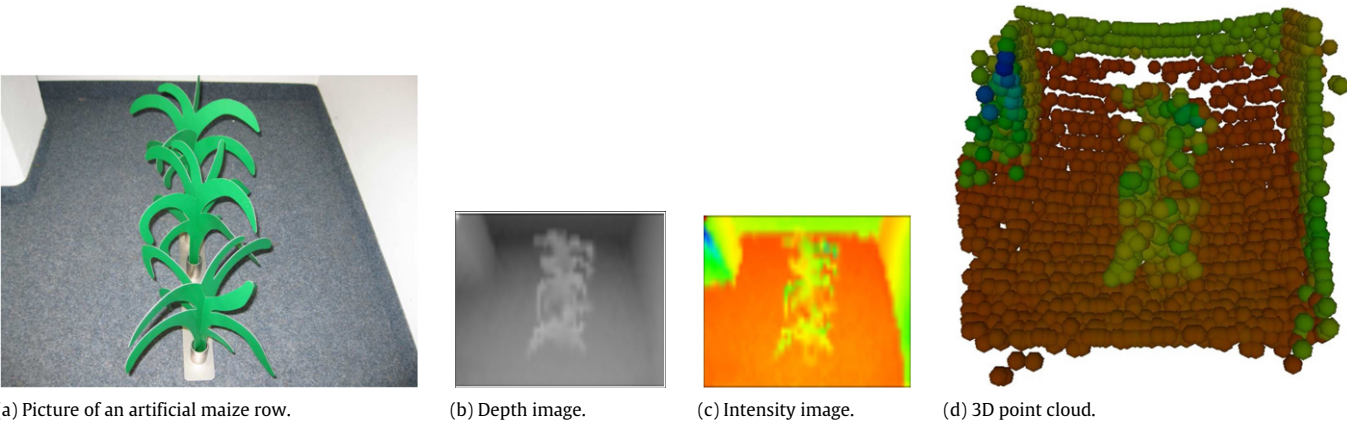


Fig. 3. The picture (a) shows an artificial maize row made of paper. The depth image (b) and the reflectance intensity image (c) below show the same view like the camera but using the FX6 sensor. The point cloud (d) on the bottom corresponds to (b) and (c).

model of the crop row and localizes the plants according to the highly precise measured pose of the robot. Therefore, the algorithm takes a 3D point cloud, the robot pose and the robot velocity as input.

Fig. 4 shows an overview over our proposed algorithm and the data flow. First, the ground is detected and a plane equation is determined. This plane equation is propagated using the robot velocity and a Kalman filter. The points which lie on the filtered ground plane are removed and the remaining points are transformed into the plane coordinate system. In the transformed

point cloud, the algorithm searches for clusters representing plants using a statistical model of the crop row. For each of these clusters the probability to be a plant is calculated. The positions of the clusters are determined according to the filtered ground plane and transformed into the world coordinate system using the robot pose. To improve the quality of the detection and the mapping the clusters are tracked over time. Finally, the positions are stored in the map, as soon as the tracked plant disappears from view. Below, the single steps and the design of the crop row model will be described in more detail.

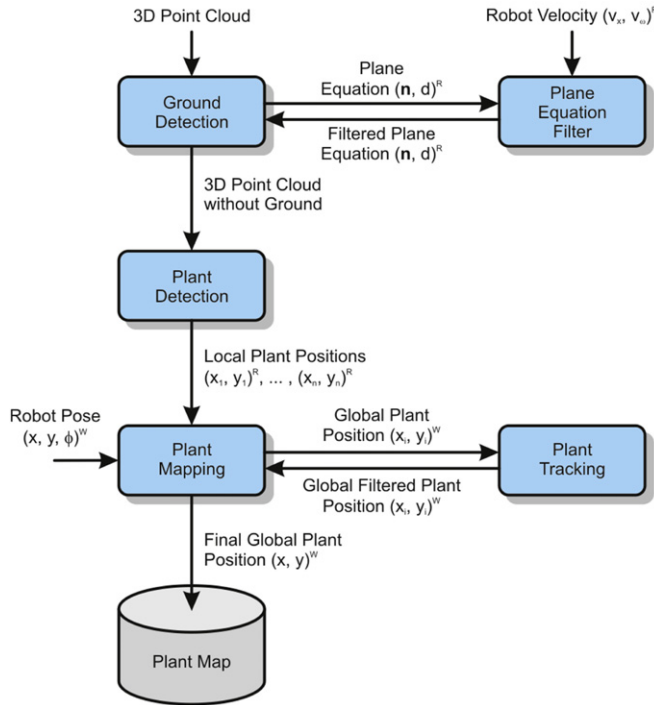


Fig. 4. The figure shows the scheme of the detection and mapping algorithm. The values labeled with a W are in world coordinates, the values labeled with an R are in robot coordinates. The algorithm takes a point cloud, detects the ground plane using a Hessian plane equation as a model, and propagates this equation using the robot velocity. Afterward, the points belonging to the ground are removed. The plant detection clusters the remaining cloud, determines the plant positions and the probability for the cluster to represent a plant. For each detected plant a filter is started which tracks the plant over time, until it disappears from view. Finally, the plant positions are stored in the plant map.

3.1. Ground detection

In our algorithm, the first step is to detect the plants in the laser scan by partitioning the points into subsets containing the ground, the plants and other structures. We utilize the fact that we know approximately the pose of the laser sensor. In our case, the scanner is mounted on a height at about 45 cm and with a tilt angle between 30° and 45° looking forward towards the ground, so we expect to see the ground plane in the laser scan (see Fig. 5(b)). Because we

are detecting the ground, we are able to determine the mounting height and angle of the sensor. As long as we are sensing the ground these parameters are variable.

To detect the ground we are fitting the Hessian plane equation

$$0 = \mathbf{n} \cdot \mathbf{x} - d, \quad (1)$$

with d , the distance to the plane and \mathbf{n} , the plane normal, as a model into the data. For the plane estimation the Random Sample Consensus (RANSAC)-Algorithm [25] is used, because it can deal with a large number of outliers. Therefore, three points from the point cloud are selected by random to set up Eq. (1). Afterward, the number of inliers, by counting the points lying on the plane within certain limits, is determined. This step is performed several times and the randomly chosen plane with the largest number of inliers is used for a refinement step. The refinement of the plane is done by using a least square fit.

Fig. 5 shows a 2D example for the plane detection using the RANSAC-Algorithm. Critical parameters for this algorithm are the number of iterations, the threshold for the maximum distance between an inlier and the plane model, and the threshold for the minimum number of inliers. These parameters were chosen using the ranging precision of the laser sensor as well as the probability for selecting a good model. Section 4 shows an experimental evaluation of these parameters.

The deviation to the plane model and the minimum number of inliers relies on the characteristic of the data. The number of required iterations can be calculated using equation

$$1 - p = (1 - w^n)^k, \quad (2)$$

with p the desired probability for selecting an good model, w the ratio between the number of inliers and the number of points in data and k the number of required iterations. Assuming that n points are required for estimating a model, w^n is the probability that all n points are inliers and $1 - w^n$ is the probability that at least one of the n points is an outlier. That probability to the power of k is the probability that the algorithm never selects a set of n points which are all inliers and this must be the same as $1 - p$. After taking the logarithm of both sides one gets

$$k = \frac{\log(1 - p)}{\log(1 - w^n)}. \quad (3)$$

We expect between 30% and 50% outlier and intend to receive a good model with a probability of 99%, so Eq. (3) gives us a number

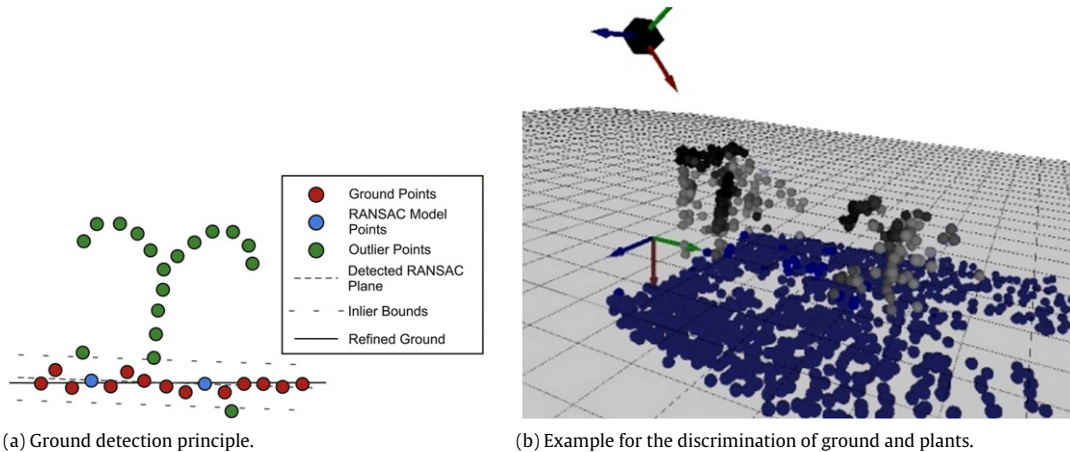


Fig. 5. The image (a) shows a synthetic 2D point cloud of a plant and a ground plane to visualize the principle of the ground detection. The blue points are the best randomly selected model with the dashed gray line as the resulting plane. The dotted gray lines display the error bounds around the model and the gray points the inliers falling into this model. The solid black line is the refined solution of the plane using least squares on the inliers. The image (b) shows the result of the ground detection for a point cloud containing ground and two plants. The points colored in blue are identified as ground, the checked ground is the resulting plane. The black box displays the sensor position and the red arrow its viewing direction. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

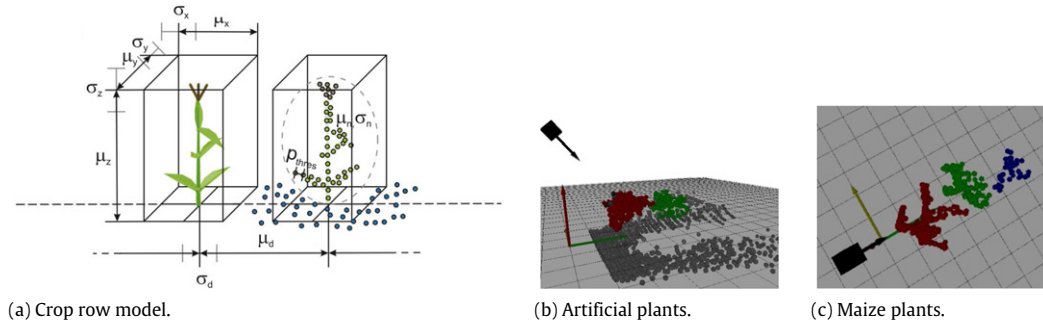


Fig. 6. Fig. (a) is showing the design of the crop row model. The model consists of the average distance μ_d between two plants, the average size of a plant represented by a bounding box (width = μ_y , height = μ_z and depth = μ_x) and the average number of points μ_n representing a plant (Points inside the ellipse) as well as the standard deviation for each of these values. The maximum deviation between two points p_{thres} is used as a parameter for the clustering. Figs. (b) and (c) show results of the clustering process. The clusters are colored in red, green and blue. The dark gray points are representing the ground plane, while the plane itself is displayed in light gray and checked (one unit = 0.1 m). The black cube displays the laser sensor position and the black arrow the viewing direction. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

between 11 and 35 iterations. The possibility of selecting only bad models is filtered by the plane equation filter, which will be described below.

After approximating the plane using the RANSAC-Algorithm, the inliers for the best model are taken to refine the plane equation using least squares. With one row for each inlier (x_i, y_i, z_i) the system of equations

$$0 = \mathbf{A} \cdot \mathbf{s} = \begin{pmatrix} x_1 & y_1 & z_1 & -1 \\ \vdots & \vdots & \vdots & \vdots \\ x_n & y_n & z_n & -1 \end{pmatrix} \cdot \begin{pmatrix} n_x \\ n_y \\ n_z \\ d \end{pmatrix} \quad (4)$$

is buildup. A least square solution for this system of equations is calculated using a singular value decomposition.

If a scan shows no ground plane or only bad models were selected, the plane detection could fail. This results in a too large number of outliers, but we are not able to define an exact threshold for that, because this threshold depends on the sensed point cloud. To get a good plane equation even in these cases and to get a smooth behavior of the plane equation, \mathbf{n} and d are propagated using the plane equation filter. Therefore, we are applying a Kalman filter. In the prediction step of this filter, the velocity of the robot is used to propagate the plane equation, and in the measurement step we are using the plane equation determined with the RANSAC-Algorithm, if a good model was selected. The noise matrix for the measurement depends on the number of inliers, the noise matrix for the prediction step on the distance traveled by the robot.

Besides the segmentation, the ground detection has several advantages and can be useful for other applications. For example one can use the information about the ground plane for calibrating the sensors position and orientation or use the ground points for terrain classification as well as for row detection or path planning algorithms.

3.2. Plant detection

While the robot is driving over a crop row more than one plant is expected in the field of view of the laser scanner, so it is not possible to take the complement of the ground as a plant. Also, other objects can occur in the field of view. Our strategy to handle this is to cluster the points which are located close to each other and consider that one cluster stands for one object. To support this clustering we are using a statistical model of a crop row containing the average distance between two plants and the expected plant size. The parameters describing a row depend on the species of the crop and the kind of the field. These parameters were experimentally determined for our tested plants

and rows. Fig. 6(a) displays the design of the model. We also have developed an approach to classify the clusters using machine learning methods and achieved discrimination rates up to 98%, which will be presented in a future publication.

For the clustering and position estimation of the plants the points are transformed into the plane coordinate system with the origin at the projection of the sensor onto the plane, the z -axis parallel to the plane normal and the x -axis pointing into the forward direction of the robot. This transformation is done, because we map the plants according to the ground plane.

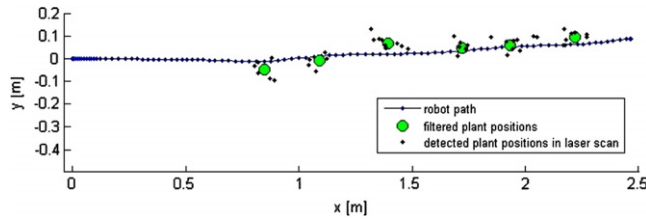
The idea behind the clustering is based on region growing in 3D. Therefore the pointcloud is organized in a k - d tree for a fast nearest neighbor search. The region growing starts at a randomly selected seed point. The neighbors of this seed point within a smaller distance than the threshold p_{thres} are removed from the pointcloud and pushed into a queue. The threshold p_{thres} depends on the row model (see Fig. 6(a)). For each point in the queue the again the nearest neighbour search is conducted, pushing also new points in the queue, until the queue is empty. If the queue is empty and at least one unclustered point is left new seed points are selected to find the next cluster. In the final step the center and the standard deviation for each cluster are calculated. Clusters with a distance smaller than the added standard deviation between the centers are merged.

After all points are clustered the positions of the plants are determined using the medians of the remaining clusters into the x and y direction. The probability that a cluster represents a plant is determined by comparing the dimensions of the bounding box around the cluster with the expected box size and the number of expected points of the cluster. For the tested plants the average values $\mu_x, \mu_y, \mu_z, \mu_n$, and the standard deviations $\sigma_x, \sigma_y, \sigma_z, \sigma_n$ were experimentally evaluated. To calculate the probability P_i for one cluster, the equations

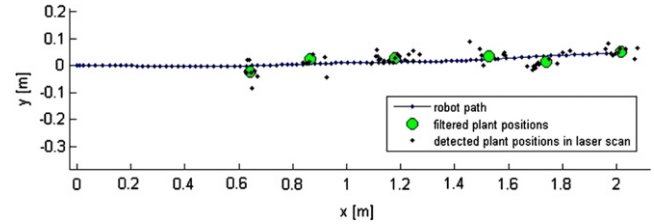
$$p_j \propto e^{-\frac{(\mu_j - \mu_j)^2}{2\sigma_j^2}} \quad \text{for } j = x, y, z, n \quad (5)$$

$$P_i = p_x \cdot p_y \cdot p_z \cdot p_n \quad (6)$$

are used, with μ_x, μ_y and μ_z , the measured bounding box dimensions and μ_n the number of measured cluster points. p_x, p_y and p_z are the resulting probabilities for the measured dimensions to correspond to a crop plant and p_n the probability for the number of cluster points corresponding to a plant. For each of these probabilities we have chosen a confidence level which implies that a cluster fitting into the plant model is rejected to be a plant at a chance of 5%, using a Z-test. Fig. 6(b) and (c) shows two example images with detected ground and the clustered plants.



(a) Mapping of artificial maize plants.



(b) Mapping of artificial plastic plants.

Fig. 7. The diagrams (a) and (b) are showing the mapping of two different crop rows, each consisting of six plants. The blue lines indicate the path traveled by the robot, the black points show the measured medians and the green circles the final positions determined by the plant tracker. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

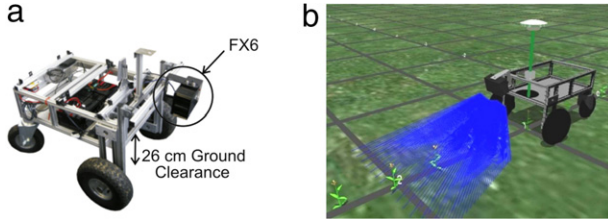


Fig. 8. The picture (a) shows the modified Volksbot RT-3 with increased ground clearance and an additional castor wheel, which enables the robot to drive over a crop row. The image (b) is a screenshot of the Volksbot and the 3D-LIDAR simulated in Gazebo.

3.3. Plant mapping

By using our robot driving over a crop row, typically two or three plants are detected at once. The positions of these plants and the associated probabilities are used as inputs for the final mapping step. To be able to store the plants in a global map, the positions have to be transformed into the world coordinate system using the knowledge of the sensors mounting position and of the pose of the robot in the world coordinate system. The height and the angle of the sensor is given by the detected ground plane automatically, so we only have to know the x and y position according to the robot origin.

To improve the localization of the plants, we are tracking them until they disappear from view. Therefore, the mapping module maintains a tracker for each plant. Each of these trackers determines the position of the plants by using the mean of the measured medians, each weighted with the corresponding probability P_i :

$$\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{\sum_i P_i} \left(\sum_i P_i \begin{pmatrix} x_i \\ y_i \end{pmatrix} \right). \quad (7)$$

To assign a plant to one tracker the plant mapping uses the crop row model. If the distance to a measured plant is smaller than a threshold of $\frac{1}{2} \cdot \mu_d$ with μ_d , the average distance between to plants, to a running tracker, the tracker which is closest to the plant is updated with the measured position and probability. If the threshold is larger than $\frac{1}{2} \cdot \mu_d$ a new tracker is started with the new position. The trackers observe the plants until they disappearing from the field of view or no new position is updated for a defined timeout. If one of these cases happen the position of the plant in the filter and the variance of the measured positions are stored in the map. Fig. 7 displays the result of this mapping process and the measured medians by means of two examples.

4. Experiments

The experimental results are partitioned into two subsections. In the first subsection we evaluated the ground detection and the

parameters for the RANSAC-Algorithm as well as the needed run-time; in the second subsection we discuss the detection rate and localization accuracy of the plant detection and mapping.

The experiments were carried out using the 3D simulation environment Gazebo [26], with artificial plants in the laboratory and on a small maize field. For the real-world experiments we have used a modified Volksbot RT-3 [27], we also simulated this robot in Gazebo with all its sensors (see Fig. 8). By using our middleware, we are able to communicate with the simulated and the real robot in the same way transparently. Table 2 gives an overview over these experimental set-ups.

4.1. Ground detection

The FX6 LIDAR provides 16 fps, so the ground plane and plant detection as well as the mapping have to be carried out in less than 62.5 ms. The mapping step is not very time consuming and the time for clustering depends linearly on the number of remaining points, after removing the ground plane. So the most expensive part of the algorithm is the ground detection.

It is necessary to find a trade-off between the detection quality and the required time. This trade-off depends on the parameters for the RANSAC-Algorithm. The most important parameter turned out to be the deviation to the plane model, which relies on the ranging standard deviation of the FX6 sensor. This parameter affects basically the quality of the plane detection. For the FX6 sensor 80 mm are specified for a deviation of $3 \cdot \sigma$ and a distance of 5 m. We evaluated this for the near distances in which we are operating and got 75 mm, approximately the same value as denoted by Nippon Signal. In the laboratory environment $1 \cdot \sigma$, and for the outdoor data a value between $1 \cdot \sigma$ and $2 \cdot \sigma$ were chosen. The experiments with these thresholds have shown that the plane can be detected reliably with a maximum deviation for d smaller than 10 mm and a maximum angle deviation of α smaller than 1° , for indoor and outdoor.

To evaluate the RANSAC-Algorithm parameters we took runs over different crop rows and measured the number of inliers, the deviation to the plane and the required time. The experiments have shown that a threshold between 20 and 40 mm is a good trade-off, larger values imply too much outliers and with a smaller threshold the plane detection rate is too small. With the selected threshold and at least 30% inliers, between 8–12 iterations are sufficient. Because of the use of the plane equation filter, these values are slightly smaller than the number of iterations calculated in Section 3.

For our plant detection experiments we have chosen the following parameters: number of iterations: 10; minimum number of inliers: 30%; deviation threshold: 30 mm. These values are also adequate for the ground detection on real field data, even on rough soil. Moreover, 10 iterations are a good trade-off between time and the detection quality. For the selected parameterization and by using no optimization by the compiler, the plane detection required approximately 15 ms on an Intel 2 GHz CPU, so we

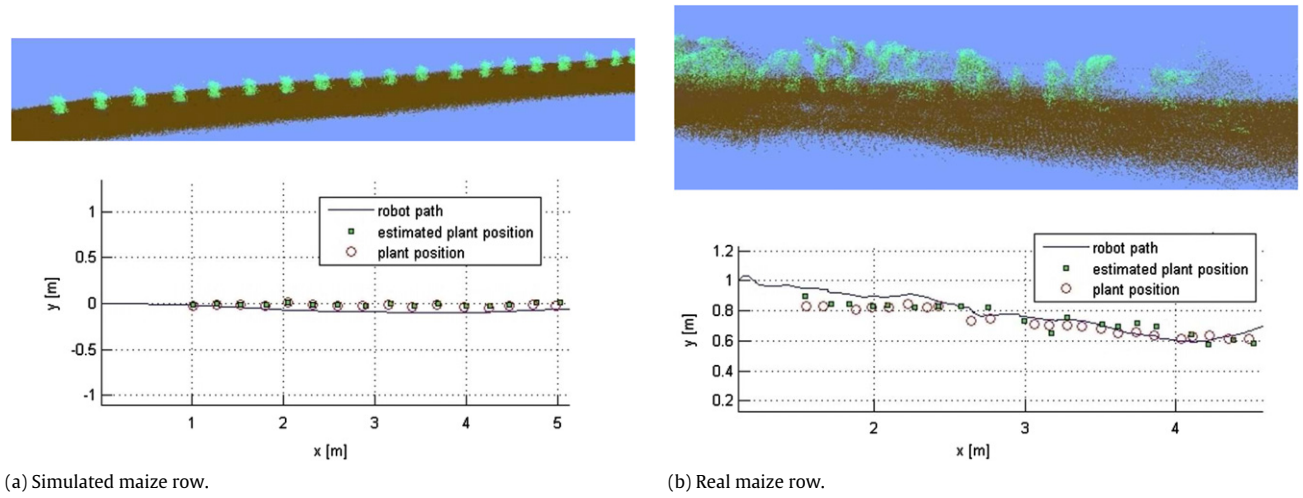






Fig. 9. Figs. (a) and (b) shows two exemplary results of the plant mapping. The point clouds were constructed by using the recorded laser scans and the robot poses. In the plant maps the blue lines indicate the paths traveled by the robot, the red circles show the ground truth positions of the plants and the green squares the estimated plant positions. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 2
Experimental field set-ups.

| | Simulation | Laboratory | | Field |
|----------------------------------|---|--|--|--|
| Plants | Simulated maize  | Paper plants  | Plastic plants  | Maize plants  |
| Plant height [m] | 0.15–0.25 | 0.2–0.25 | ca. 0.2 | 0.125–0.25 |
| Plant displacement [m] | 0.1–0.25 | ca. 0.2 | ca. 0.2 | 0.08–0.15 |
| Row length [m] | ca. 10 | ca. 1.4 | ca. 1.4 | ca. 8 |
| Average number of plants per row | 40 | 6 | 6 | 45 |

have about 45 ms left for the plant detection and mapping. Using 12 iterations the whole algorithm required about 27 ms with a standard deviation of about 2.1 ms, for 25 iterations the we got 47 ms with a standard deviation of about 2.4 ms.

4.2. Plant detection and mapping

For the plant detection algorithm a maximum clustering threshold pthres of about 25 mm between two neighbors has shown to provide reliable plant detection within a smaller time period than the available 35–45 ms. The algorithm even works well if the plants grow close to each other. The detection rate for a plant in the simulation environment is about 99%, in the indoor and outdoor experiments we got about 60%–70%. Using the plant tracker this values are good enough for the mapping (see Table 3).

To evaluate the plant mapping we need to know the exact position of the plants. In the simulation environment we have the exact positions of the plants by definition, in the laboratory we measured the positions and the starting pose of the robot with a yard stick. Outdoors we got the start position of robot using the GPS. But to get the plant positions of real plants is a much more complex job, compared to the both other scenarios. Therefore, we used our robot and the RTK-GPS. We drove the robots origin over each plant and recorded the position. Towards the direction of the row this measurement is precise (about 1–2 cm), but vertically to the row the measured position is not that precise. Because of that, we determined the accuracy of our mapping algorithm by using only the deviation into the row direction.

Table 3
Detection results.

| | Simulation | Laboratory | Field |
|--|------------|------------|----------|
| Average accuracy of robot pose [m] | 0.001 | ca. 0.01 | ca. 0.02 |
| Average detection rate [%] | ca. 99 | ca. 70 | ca. 60 |
| Average detection rate using tracker [%] | >99 | >90 | 80–90 |
| Average position accuracy [m] | 0.005 | ca. 0.03 | ca. 0.03 |
| Std position accuracy [m] | 0.01 | 0.03 | 0.035 |

The localization accuracy of the plants depends strongly on the accuracy of the robot pose. In the simulation environment and outdoors we determine this pose with an RTK-GPS, the odometer and an IMU, indoors we just use the odometer values. Table 3 gives an overview over the average accuracy of the robot pose, the average detection rate and position accuracy for our different experimental set-ups. Fig. 9 shows two exemplary maps and point clouds of a simulated and a real crop row. The detection rate and localization accuracy using the simulation environment is very good, because the simulated crop rows fits nearly perfect to the crop row model. For the real world experiments the detection fails if the deviation between the plants diers too much from the model. For example see plant number 7 and 14 in Fig. 9(b). Further, the position error for the real world experiments into the y-axis of the robot is relatively heigh compared to the error into the x-axis. This is because of the localization heading error. In the meantime the localization algorithm was improved, but since that new experiments for plant detection have not been carried out. Another error source for the plant localization is biologically aected. The ground truth positions of the plants were measured on the ground plane, but not each plant is really growing upright.

5. Conclusions and future works

In this work we evaluated the state-of-the-art for environment sensing in the domain of agricultural robotics and discussed the advantages of 3D sensors in such applications. 3D laser sensors seems to be the most promising and reliable systems for autonomous agricultural robots to sense the environment, with the advantages that they even work under various weather conditions as well as in foggy and dusty environments, which are not uncommon in agricultural applications.

At present, the 3D FX6 LIDAR used by us, has a relatively small resolution. Nevertheless, we have shown that by using the algorithm we have developed, an agricultural robot is able to detect reliably single plants in crop rows in real-time. Now, as the positions of the single plants can be determined the task of repeating phenotyping of individual plants at different days could be performed. Individual plant detection offers further possibilities of single plant care like spraying, fertilization or weeding which leads to a reduction of costs, fertilizers and pesticides, and to a dramatic improvement in domain of agricultural operations.

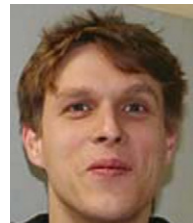
In the future we intend to advance the plant detection using an improved model of the crop row as well as machine learning methods to discriminate several kinds of plants and to learn new plant species autonomously. These machine learning methods also may improve the clustering algorithm, by learning criteria if two clusters should be connected or if a cluster should be split off. Further, we will evaluate how good the plants can be retrieved and if we are able to improve the localization by using the plants as natural landmarks. Besides the single plant detection we are working at row detection algorithms, to be able to navigate independently over the rows, without caring about the individual plants.

References

- [1] B.J. Legg, J.V. Stafford, Precision agriculture – new technologies, in: The Brighton Conference – Pest and Diseases, 1998, pp. 1143–1150.
- [2] Y. Edan, S. Han, N. Kondo, Automation in agriculture, in: Handbook of Automation, Springer, 2009, pp. 1095–1128.
- [3] A. Ruckelshausen, P. Biber, M. Dorna, H. Gremmes, R. Klose, A. Linz, F. Rahe, R. Resch, M. Thiel, D. Trautz, U. Weiss, Bonirob – an autonomous field robot platform for individual plant phenotyping, in: Joint International Agricultural Conference, 2009, p. 464.
- [4] Nippon-Signal, January 2009. FX6. <http://www.signal.co.jp/>.
- [5] H. Auernhammer, T. Muhr, Gps in a basic rule for environment protection in agriculture, in: Symposium Automated Agriculture in the 21st Century. St. Joseph, USA, 1991, pp. 394–402.
- [6] J.N. Wilson, Guidance of agricultural vehicles – a historical perspective, Computers and Electronics in Agriculture 25 (2000) 3–9.
- [7] B. Åstrand, A.-J. Baerveldt, A vision based row-following system for agricultural field machinery, Mechatronics 15 (2) (2004) 251–269.
- [8] A. Tellaache, X.P. Burgos-Artizzu, G. Pajares, A.R., On combining support vector machines and fuzzy k-means in vision-based precision agriculture, in: World Academy of Science, Engineering and Technology. Vol. 22, July 2007, pp. 188–193.
- [9] J. Billingsley, Some agricultural applications of machine vision, International Conference on Robotics and Automation 1 (2008) 10–15.
- [10] L. Tang, J. Jin, Corn plant sensing with active and passive stereo vision, International Conference on Robotics and Automation 1 (2008) 31–37.
- [11] F. Rovira-Más, Q. Zhang, M. Kise, J.F. Reid, Agricultural 3d maps with stereovision. Position, Location, And Navigation Symposium, 2006, pp. 1045–1053.
- [12] J. Wei, F. Rovira-Mas, J. Reid, S. Han, Obstacle detection using stereo vision to enhance safety of autonomous machines, Transactions of the ASAE 48 (2005) 2389–2397.
- [13] T. Satow, K. Matsuda, S.B. Ming, K. Hironaka, D.L.S. Tan, Development of laser crop row sensors for automatic guidance system of implements. Automation Technology for Off-Road Equipment, 2004.
- [14] F. Fender, M. Hanneken, S.I. der Stroth, A. Kielhorn, A. Linz, A. Ruckelshausen, Sensor Fusion Meets GPS: Individual plant detection, in: Proceedings of CIGR EurAgEng/VDI-MEG, 2006.
- [15] J. Weingarten, G. Gruener, R. Siegwart, A state-of-the-art 3d sensor for robot navigation, in: International Conference on Intelligent Robots and Systems, 2004, pp. 2155–2160.
- [16] S. May, B. Werner, H. Surmann, K. Pervölz, 3d time-of-flight cameras for mobile robotics, in: International Conference on Intelligent Robots and Systems, 2006, pp. 790–795.
- [17] P. Einramhof, S. Olufs, M. Vincze, Experimental evaluation of state of the art 3d-sensors for mobile robot navigation, in: 31st Austrian Association for Pattern Recognition Workshop, 2007.
- [18] T. Moeller, H. Kraft, J. Frey, M. Albrecht, R. Lange, Robust 3d measurement with pmd sensors, Tech. rep., PMDTec, 2005.
- [19] R. Klose, J. Penlington, A. Ruckelshausen, Usability study of 3d time-of-flight cameras for automatic plant phenotyping. Bornimer Agrartechnische Berichte 69, image Analysis for Agricultural Products and Processes, 2009.
- [20] H. Surmann, K. Lingemann, A. Nüchter, J. Hertzberg, A 3d laser range finder for autonomous mobile robots, 2001.
- [21] S. Thrun, W. Burgard, D. Fox, A real-time algorithm for mobile robot mapping with applications to multi-robot and 3d mapping, in: International Conference on Robotics and Automation, 2000, pp. 321–328.
- [22] P. Allen, I. Stamos, A. Gueorguiev, E. Gold, P. Blaer, Avenue: Automated site modelling in urban environments, in: Third International Conference on 3D Digital Imaging and Modeling, 2001, pp. 357–364.
- [23] V. Sequeira, K.C. Ng, E. Wolfart, J.G. Goncalves, D.C. Hogg, Automated 3d reconstruction of interiors with multiple scan views, Conf. & Publ. on Photonics, Optics, & Imaging 3641 (1998) 106–117.
- [24] M. Hebert, M. Deans, D. Huber, B. Nabbe, N. Vandapel, Progress in 3-d mapping and localization, in: International Symposium on Intelligent Robotic Systems, 2001, pp. 145–154.
- [25] M.A. Fischler, R. Bolles, Random sample consensus. a paradigm for model fitting with appchahons to image analysm and automated cartography, in: Image Understanding Workshop (College Park, Md., Apr 1980), 1980, pp. 71–88.
- [26] N. Koenig, Gazebo. May 2009. <http://playerstage.sourceforge.net/index.php>.
- [27] H. Surmann, Volksbot, March 2009. <http://www.volksbot.de/>, fraunhofer Institut Intelligente Analyse- und Informationssysteme.



Ulrich Weiss received his diploma degree in Computer Science from the University of Tübingen and is a Ph.D. student in the Corporate Sector Research and Advance Engineering at the Robert Bosch GmbH. His research is on localization and navigation of autonomous agricultural robots.



Peter Biber is a research engineer and a project manager at the Corporate Sector Research and Advance Engineering of the Robert Bosch GmbH in Schwieberdingen, Germany. He earned his diploma degree (Dipl.-Inf.) and his doctoral degree (Dr.rer. nat) from the University of Tübingen, Germany. During his doctoral studies he was a research assistant at the department of Graphical-Interactive Systems at the Wilhelm Schickard Institute for Computer Science at the University of Tübingen and a visiting researcher at the Center for Applied Autonomous Sensor Systems (AASS) at Örebro University, Sweden. His research interests include robotics, computer vision and computer graphics, especially topics like localization, sensor fusion, sensor data interpretation and 2D/3D reconstruction and mapping. He is the author and the co-author of more than 25 papers in peer-reviewed journals and international conference proceedings and holds more than 10 patents.