Dissertation

Crop plant reconstruction and feature extraction based on 3-D vision

Submitted in fulfilment of the requirements for the degree

"Doktor der Agrarwissenschaften"

(Dr.sc.agr. in Agricultural Science)

To the

Faculty of Agricultural Sciences

Presented by: **M.Sc. Manuel Vázquez Arellano**

Born in: Mexico

Year of publication: 2019

This thesis was accepted as a doctoral thesis (Dissertation) in fulfilment of the regulations to acquire the degree "Doktor der Agrarwissenschaften" (Dr.sc.agr. in Agricultural Sciences) by the Faculty of Agricultural Sciences at University of Hohenheim on 05.11.2018.

Date of oral examination: 15. January 2019

Examination Committee

Head of the Committee:	Prof. Dr. Andrea Knierim
Supervisor and 1st Examiner:	Prof. Dr. Hans W. Griepentrog
2nd Examiner:	Prof. Dr. Constantino Valero Ubierna
3rd Examiner:	Prof. Dr. Stefan Böttinger

Self-publishing:	Manuel Vázquez Arellano
Source of supply:	Institute of Agricultural Engineering of
	the University of Hohenheim
	Garbenstraße 9, 70599 Stuttgart

Summary

3-D imaging is increasingly affordable and offers new possibilities for a more efficient agricultural practice with the use of highly advances technological devices. Some reasons contributing to this possibility include the continuous increase in computer processing power, the decrease in cost and size of electronics, the increase in solid state illumination efficiency and the need for greater knowledge and care of the individual crops. The implementation of 3-D imaging systems in agriculture is impeded by the economic justification of using expensive devices for producing relative low-cost seasonal products. However, this may no longer be true since low-cost 3-D sensors, such as the one used in this work, with advance technical capabilities are already available.

The aim of this cumulative dissertation was to develop new methodologies to reconstruct the 3-D shape of agricultural environment in order to recognized and quantitatively describe structures, in this case: maize plants, for agricultural applications such as plant breeding and precision farming. To fulfil this aim a comprehensive review of the 3-D imaging systems in agricultural applications was done to select a sensor that was affordable and has not been fully investigated in agricultural environments. A low-cost TOF sensor was selected to obtain 3-D data of maize plants and a new adaptive methodology was proposed for point cloud rigid registration and stitching. The resulting maize 3-D point clouds were highly dense and generated in a cost-effective manner. 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. The generated point cloud was used to obtain information about the plant parameters (stem position and plant height) in order to quantitatively describe the plant. The resulting plant stem positions were estimated with an average mean error and standard deviation of 27 mm and 14 mm, respectively. Additionally, meaningful information about the plant height profile was also provided, with an average overall mean error of 8.7 mm. Since the maize plants considered in this research were highly heterogeneous in height, some of them had folded leaves and were planted with standard deviations that emulate the real performance of a seeder; it can be said that the experimental maize setup was a difficult scenario. Therefore, a better performance, for both, plant stem position and height estimation could be expected for a maize field in better conditions. Finally, having a 3-D reconstruction of the maize plants using a cost-effective sensor, mounted on a small electric-motor-driven robotic platform, means that the cost (either economic, energetic or time) of generating every point in the point cloud is greatly reduced compared with previous researches.

Zusammenfassung

Die 3D-Bilderfassung ist zunehmend kostengünstiger geworden und bietet neue Möglichkeiten für eine effizientere landwirtschaftliche Praxis durch den Einsatz hochentwickelter technologischer Geräte. Einige Gründe, die diese ermöglichen, ist das kontinuierliche Wachstum der Computerrechenleistung, die Kostenreduktion und Miniaturisierung der Elektronik, die erhöhte Beleuchtungseffizienz und die Notwendigkeit einer besseren Kenntnis und Pflege der einzelnen Pflanzen. Die Implementierung von 3-D-Sensoren in der Landwirtschaft wird durch die wirtschaftliche Rechtfertigung der Verwendung teurer Geräte zur Herstellung von kostengünstigen Saisonprodukten verhindert. Dies ist jedoch nicht mehr länger der Fall, da kostengünstige 3-D-Sensoren, bereits verfügbar sind. Wie derjenige dier in dieser Arbeit verwendet wurde.

Das Ziel dieser kumulativen Dissertation war, neue Methoden für die Visualisierung die 3-D-Form der landwirtschaftlichen Umgebung zu entwickeln, um Strukturen quantitativ zu beschreiben: in diesem Fall Maispflanzen für landwirtschaftliche Anwendungen wie Pflanzenzüchtung und „Precision Farming" zu erkennen. Damit dieses Ziel erreicht wird, wurde eine umfassende Überprüfung der 3D-Bildgebungssysteme in landwirtschaftlichen Anwendungen durchgeführt, um einen Sensor auszuwählen, der erschwinglich und in landwirtschaftlichen Umgebungen noch nicht ausgiebig getestet wurde. Ein kostengünstiger TOF-Sensor wurde ausgewählt, um 3-D-Daten von Maispflanzen zu erhalten und eine neue adaptive Methodik wurde für die Ausrichtung von Punktwolken vorgeschlagen. Die resultierenden Mais-3-D-Punktwolken hatten eine hohe Punktedichte und waren in einer kosteneffektiven Weise erzeugt worden. Die Validierung der Methodik zeigte, dass die Pflanzen mit hoher Genauigkeit rekonstruiert wurden und die qualitative Analyse die visuelle Variabilität der Pflanzen in Abhängigkeit der 3-D-Perspektive zeigte. Die erzeugte Punktwolke wurde verwendet, um Informationen über die Pflanzenparameter (Stammposition und Pflanzenhöhe) zu erhalten, die die Pflanze quantitativ beschreibt. Die resultierenden Pflanzenstammpositionen wurden mit einem durchschnittlichen mittleren Fehler und einer Standardabweichung von 27 mm bzw. 14 mm berechnet. Zusätzlich wurden aussagekräftige Informationen zum Pflanzenhöhenprofil mit einem durchschnittlichen Gesamtfehler von 8,7 mm bereitgestellt. Da die untersuchten Maispflanzen in der Höhe sehr heterogen waren, hatten einige von ihnen gefaltete Blätter und wurden mit Standardabweichungen gepflanzt, die die tatsächliche Genauigkeit einer Sämaschine nachahmen. Man kann sagen, dass der experimentelle Versuch ein schwieriges Szenario war. Daher könnte für ein Maisfeld unter besseren Bedingungen eine besseres Resultat sowohl für die Pflanzenstammposition als auch

für die Höhenschätzung erwartet werden. Schließlich bedeutet eine 3D-Rekonstruktion der Maispflanzen mit einem kostengünstigen Sensor, der auf einer kleinen elektrischen, motorbetriebenen Roboterplattform montiert ist, dass die Kosten (entweder wirtschaftlich, energetisch oder zeitlich) für die Erzeugung jedes Punktes in den Punktwolken im Vergleich zu früheren Untersuchungen stark reduziert werden.

Table of contents

List of figures

List of tables

List of abbreviations

CCD   Charge-coupled device

CTC   Consumer time-of-flight camera

CTS   Consumer triangulation sensors

DBSCAN   Density-based spatial clustering of applications with noise

GNSS   Global navigation satellite system

ICP   Iterative Closest Point

LAI   Leaf area index

LAR   Least absolute residuals

LIDAR   Light detection and ranging

NIR   Near infrared

RANSAC   Random Sample Consensus

TOF   Time-of-flight

UAV   Unmanned aerial vehicle

1-D   One dimensional

2-D   Two dimensional

3-D   Three dimensional

1 General introduction

3-D imaging systems are a key technological component for robotics. The reason behind it is that 3-D perception provides more awareness to the system than 2-D does. Also, it allows greater accuracy to judge distances, therefore, it is more robust for tasks that require high precision at the cost of extra processing. This fact applies not only in 3-D imaging systems but also in nature with complex organisms such as primates and humans. It is considered that the development of the robust forward-pointing stereoscopic vision, found in primates and humans, was the main driver of their large brain size due to the increased amount of processing that it required (Bradbury and Vehrencamp, 2011). Therefore, 3-D imaging systems should be considered in complex agricultural tasks that require high accuracy in structure recognition. 3-D imaging systems can provide useful information for precision agriculture, increased perception for agricultural robotics and better screening workflows for plant breeding. Since 3-D information can describe the complex plant architecture and its changes in time with a unique non-invasive approach, new applications are foreseen in the near future.

Among the different principles for measuring depth, and thus obtaining 3-D data, time-of-flight is increasingly appealing due to some inherent technical characteristics that make it very promising.

1.1 Problem statement

Economically affordable sensors are a key factor for the implementation of principles of precision farming as well as of highly automated systems like robots. Especially the use of 3-D imaging systems in agriculture show high potential and if being affordable they would be more appealing and accessible. There is the need of cost-effective 3-D systems to increase their viability in agricultural applications, however, there are still some questions regarding their technical capabilities, advantages and their absence in commercial solutions, among others.

1.2 The sensor (Kinect v2 camera)

In general, TOF cameras offer a suitable solution for depth measurement due to their increased improvement in technical characteristics (i.e. sensing range and accuracy) (Lange et al., 1999) and their decreasing cost per pixel (Frey, 2010). However, there are other reasons such as:

− No shadow effects since illumination and observation directions are collinear.

− No mechanical moving parts that make it more robust for the rough agricultural environments.

− Scene illumination is provided by the camera (active illumination), and is not dependent on external illumination.

− Since the wavelength of the measuring light rays are in the near infrared range, some cameras offer infrared stream output that could be used for night farming. Additionally, the plant reflectance signature in the near infrared is high, in the so called *near infrared plateau* (Govender et al., 2009), therefore a good quality depth measurements are possible at night.

In this research, the data acquisition was done using a TOF camera that was manufacture for the consumer electronics market, specifically as a gaming technology device for human-computer interaction. The first Kinect (Kinect v1) was launched on November , 2010 and sold 8 million units in the first 60 days entering the Guinness World Records as the "fastest selling consumer electronics device in history" (Ramos, 2012). The Kinect v2, used in this research, was launched on November 2013 but since then, there were several concerns in their main market such as reliability, operation conditions (space requirements) and privacy concerns regarding mass data surveillance.

There are several reports that state that Microsoft is discontinuing the Kinect v2 sensor and its adaptor. Additionally, at the time of writing, the sensor is out of stock in the official Microsoft store, and since the main market of the Kinect v2 is the consumer electronics, its commercial performance in this sector determines its price and availability. Currently, virtual reality glasses are taking over the human-machine interaction that was the main task of the Kinect sensor, therefore, there are several reasons to question its future commercial availability.

1.3   Aim and objectives

The aim of this work was to develop new methodologies for reconstructing agricultural structures such as the maize plant morphology by assembling (or stitching) 3-D images and extract plant features. In order to accomplish this aim, the following objectives should be fulfilled:

− To review the state-of-the-art of 3-D imaging techniques with a focus on agricultural applications

− Validation of generated point clouds based on 3-D image registration and stitching methodologies.

– Feature extraction and validation from point clouds such as individual plant segmentation, plant height, plant or stem position and longitudinal height profile of plants.

1.4 Appended papers

This dissertation is divided in four parts based on the following papers:

Part I   Vázquez-Arellano, M., Griepentrog, H. W., Reiser, D., & Paraforos, D. S. (2016). 3-D Imaging Systems for Agricultural Applications — A Review. *Sensors*, *16*(5), 618. doi:10.3390/s16050618

Part II   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.

Part III   Vázquez-Arellano, M., Reiser, D., Paraforos, D. S., Garrido-Izard, M., Burce, M. E. C., & Griepentrog, H. W. (2018). 3-D reconstruction of maize plants using a time-of-flight camera. *Computers and Electronics in Agriculture*, *145*, 235–247. doi:10.1016/j.compag.2018.01.002

Part IV   Vázquez-Arellano, M., Paraforos, D. S., Reiser, D., Garrido-Izard, Griepentrog, H. W. (2018). Determination of stem position and height of reconstructed maize plants using a time-of-flight camera. *Computers and Electronics in Agriculture*, 154, 276-288. doi:10.1016/j.compag.2018.09.006

This thesis is structured in three parts. In Part I and extensive review of 3-D imaging systems in agriculture was done. The review considered 3-D imaging systems based on the three different basic principles for 3-D image generation: triangulation, TOF and interferometry. The review was focused on three agricultural applications: vehicle navigation, crop husbandry and animal husbandry. This section concludes with a future outlook of the technology and the identification of research gaps that need to be addressed.

In Part II the consumer TOF camera used in this research was tested in different agricultural environment, in order to evaluate the effect of varied lighting conditions on the different output streams of the camera: RGB, infrared and depth images.

In Part III a methodology for 3-D imaging reconstruction of maize was developed by using a low-cost TOF camera. This publication proposes a pipeline for 3-D image registration and stitching. The aim was to reconstruct maize plants, inside a greenhouse, and evaluate its precision.

In Part IV a methodology for estimating the stem position of reconstructed maize plants was developed. The maize plant point clouds considered in this publication were taken from the publication in Part II. Additionally, another methodology for maize plant height estimation and height profile was developed using the estimated stem positions.

1.5 References

Bradbury, J., Vehrencamp, S., 2011. Principles of Animal Communication, 2nd ed. Sinauer Associates Inc.

Frey, V., 2010. PMD Cameras for Automotive & Outdoor Applications. 3.

Govender, M., Dye, P., Weiersbye, I.., Witokwski, E.T.., Ahmed, F., 2009. Review of commonly used remote sensing and ground based technologies to measure plant water stress. Water SA 35, 741–752. doi:10.4314/wsa.v35i5.49201

Lange, R., Seitz, P., Biber, A., Schwarte, R., Inv, N., Siegen, D.-, 1999. Time-of-flight range imaging with a custom solid-state image sensor. Laser Metrol. Insp. 3823, 1–12.

Ramos, E., 2012. Kinect Basics. Arduino Kinect Proj. 23–34. doi:10.1007/978-1-4302-4168-3_2

2 Part I: 3-D imaging systems for agricultural applications – A review

Abstract

Efficiency increase of resources through automation of agriculture requires more information about the production process, as well as process and machinery status. Sensors are necessary for monitoring the status and condition of production by recognizing the surrounding structures such as objects, field structures, natural or artificial markers, and obstacles. Currently, 3-D sensors are economically affordable and technologically advanced to a great extent, so a breakthrough is already possible if enough research projects are commercialized. The aim of this review paper is to investigate the state-of-the-art of 3-D vision systems in agriculture, and the role and value that only 3-D data can have to provide information about environmental structures based on the recent progress in optical 3-D sensors. The structure of this research consists of an overview of the different optical 3-D vision techniques, based on the basic principles. Afterwards, their application in agriculture are reviewed. The main focus lays on vehicle navigation, and crop and animal husbandry. The depth dimension brought by 3-D sensors provides key information that greatly facilitates the implementation of automation and robotics in agriculture.

Keywords: 3-D sensors; optical triangulation; time-of-flight; interferometry; agricultural automation; agricultural robotics

2.1 Introduction

Sustainable strategies are in demand in agriculture due to the urgent need to increase resource efficiency in crop production. Agricultural mechanization and intensification have greatly contributed to the development of a food production system able to provide food, feed, fibre and even fuel for the world's population. Unfortunately, large amount of resources like: fuel, water, herbicides, pesticides, and fertilizers have been intensely employed in this, resulting in a current environmentally unsustainable situation due to the low resource use efficiency (UNEP, 2010).

Agricultural automation and robotics can play a significant role in society to meet its future food production needs (Bergerman et al., 2013). These technologies have already lowered production costs, reduced the intensive manual labour, raised the quality of farm products and improved environmental control (Edan et al., 2009). 1-D (Joergensen, 2002) and 2-D vision systems

have been an integral part of the successful implementation of agricultural automation and ro-botics in the food production chain. It is believed that the machine vision technology is at an inflection point, moving into a 3-D approach, driven by the improved technology and lower device prices in the consumer market (Eddershaw, 2014). In the last decade, the number of publications related to agricultural 3-D vision systems has been growing fast. Some reasons contributing to this tendency include the continuous increase in computer processing power, the decrease in cost and size of electronics, the increase in solid state illumination efficiency, the unique non-contact and non-destructive properties of machine vision technology, and the need for greater knowledge and care of the individual crops. The implementation of 3-D imag-ing systems in agriculture is impeded by the economic justification of using expensive devices for producing relative low-cost seasonal products. However, this may no longer be true since actuators, not vision sensors, are now among the most expensive components in automated and even robotic systems.

In 3-D vision systems, a single image of a scene contains huge amounts of information where recovery of depth information is complex. The depth of information is lost in the projection from the 3-D world to the 2-D imaging surface (Ma et al., 2004). Due to the extra dimension, a 3-D image increases in the amount of data that needs to be handled, and as a consequence increases the significance of the 3-D image generation techniques. These techniques are vital for handling the extraction of depth information. This is particularly evident in the triangulation techniques where algorithms get a collection of images of a scene and extract the 3-D information out of it (Bellmann et al., 2007). Figure 2-1 distinguishes between 3-D generation and image processing techniques. 3-D image generation techniques are critical for producing useful 3-D raw data. On the other hand, 3-D image processing techniques are important for processing information out of the 3-D image. They are not considered in the present study since they are highly dependent on the 3-D raw data quality.
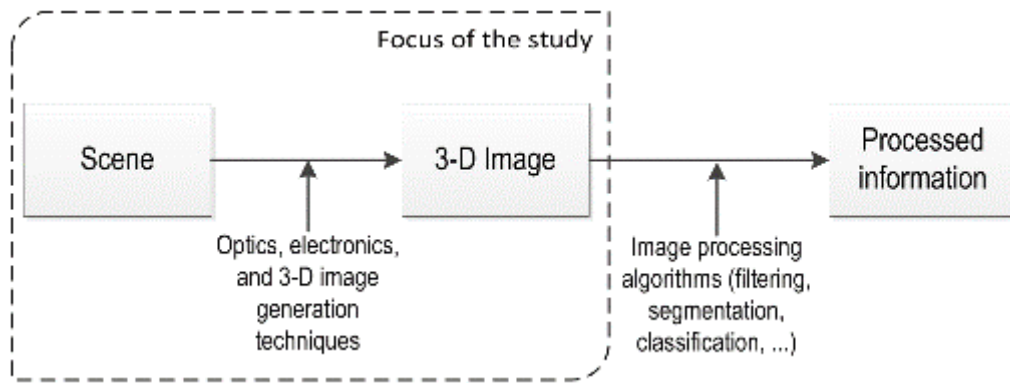
Figure 2-1:   The 3-D image generation techniques are critical for generating robust raw data
              for useful information extraction.

Review articles are frequently needed in areas featuring a rapidly growing number of research papers. There are already some reviews detailing the different techniques for 3-D image acquisition (Blais, 2004; Jarvis, 1983) and even some reviews of 2-D together with some 3-D vision systems in agricultural applications (Grift, 2008; McCarthy et al., 2010). However, there has been no comprehensive review so far that provides an insight into the achievements and potential of 3-D vision systems in agricultural applications.

The aim of this review paper is thus to investigate the state-of-the-art of 3-D vision systems in agriculture, and the role, value, and advantages that only 3-D data can have to provide information about the surrounding structures such as objects, field structures, natural or artificial markers, and obstacles based on the recent progress in optical 3-D sensors. The structure of this paper consists of an overview of the different optical 3-D vision techniques based on the basic principles. Afterwards, their application in agriculture is reviewed. The review specifically focuses on vehicle navigation and crop and animal husbandry.

2.2   3-D vision techniques

A 3-D image is a large collection of distance measurements from a known reference coordinate system to surface points on the objects scene (Besl, 1988). Depending on the context, a 3-D image is also known as range image, depth map, depth image, or 2.5-D image. In surveying, terms like Digital Terrain Map (DTM), Digital Elevation Model (DEM), and Digital Surface Model (DSM) are commonly used. Several types of spectral waves like light, ultrasound and microwaves can be used to retrieve depth information. 3-D data acquisition with optical techniques is favoured over other alternatives since optical systems allow fast 3-D shape measurement acquisition, high lateral resolution and safety standards compliance (Büttgen et al., 2005).

Several classifications have been proposed based on common characteristics, but the one based on the basic principles described by Schwarte et al. (1999) is widely accepted and provides a detailed insight of the varied possibilities in a well-organized and hierarchical structure. This classification provides more information about the principle behind the construction of a 3-D image, and a more detailed description about the different techniques and applications can be found in the Handbook of computer vision and applications (Jähne et al., 1999). These basic principles for optical depth measurement are triangulation, TOF, and interferometry.

2.2.1   Triangulation

Triangulation is a geometrical calculation where the target is one point of a triangle and the other two points are known parts of the measurement system. Measuring the triangle's angles or baseline, the distance to the target can be determined (Lange, 2000). Triangulation is the most commonly used principle for depth measurement. Figure 2-2 shows two typical examples of triangulation based techniques using active and passive illumination.
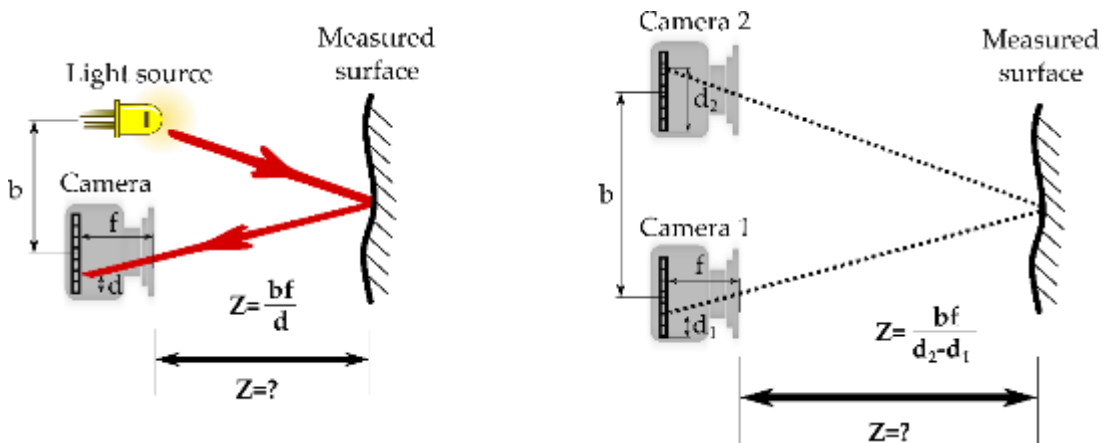
Figure 2-2:   Schematic representation of light beam (left) and stereo vision (right) triangulation. "Z", depth; "b", baseline length; "d", position of the incoming light beam on the image sensor; and "f", focal length.

Triangulation is divided into a variety of techniques based on visual cues to infer depth. Table 2-1 lists these techniques under the different triangulation approaches.

Triangulation Approach	Visual Cue	3-D Image Generation Techniques
Digital photogrammetry	Stereopsis	Stereo vision (Scharstein and Szeliski, 2002)
		Multi-view stereo (Seitz et al., 2006)
		Multiple-baseline stereo (Okutomi and Kanade, 1993)
	Motion	Structure-from-motion (Westoby et al., 2012)
		Shape-from-zooming (Lavest et al., 1993)
		Optical flow (Sun, 2002)
	Silhouette	Shape-from-silhouette (Cheung et al., 2005)
		Shape-from-photoconsistency (Kutulakos and Seitz, 2000)
		Shape-from-shadow (Savarese, 2005)
Structured light	Texture	Shape-from-texture (Lobay and Forsyth, 2006) Shape-from-structured light (Salvi et al., 2010)
Shading	Shading	Shape-from-shading (Horn, 1970)
		Photometric stereo (Woodham, 1980)
Focus	Focus	Shape-from-focus (Nayar and Nakagawa, 1994)
		Shape-from-defocus (Favaro and Soatto, 2005)
Theodolite	Stereopsis	Trigonometry (Tiziani, 1989)

Table 2-1: List of some triangulation techniques for 3-D image generation found in the literature for different visual cues.

There are digital photogrammetry (passive), structured light (active), shading, focus and theodolite measuring techniques. The Xtion Pro sensor (ASUS, Taipei, Taiwan) and the first generation Kinect™ (Kinect v1, Microsoft Corp., Albuquerque, New Mexico, USA) are examples of consumer triangulation sensors (CTSs) based on the structured light volume technique that use a pseudo random pattern to retrieve depth.

2.2.2   TOF

TOF sensors measure depth using the known speed of light and its time of flight directly or indirectly (Figure 2-3). Sensors such as LIDARs, Flash LIDARs, 360° 3-D LIDARs, and TOF cameras belong to this category. TOF depth measurement principles can be divided into pulse modulation, continuous wave modulation, and pseudo-noise modulation.
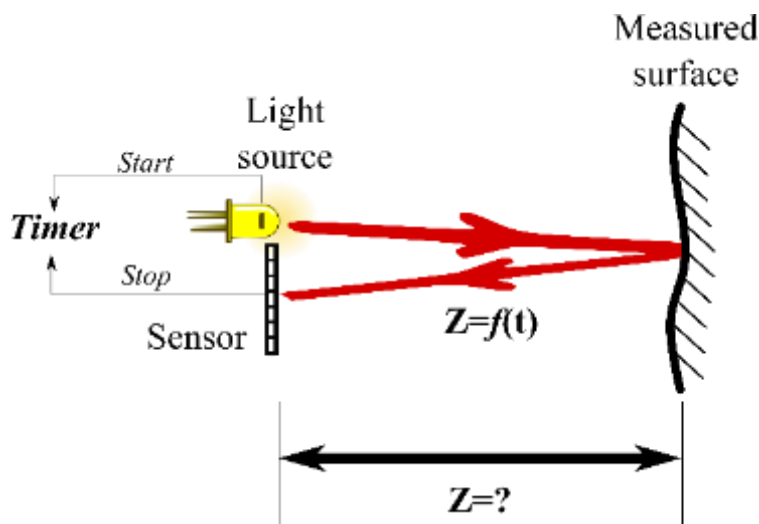
Figure 2-3:   Schematic representation of the basic principle of time-of-flight measurement, where distance "Z" is dependent on the time "t" that takes a light pulse to travel forth and back.

TOF cameras are available since the last decade and are increasingly being used in agricultural applications. They are known in the literature for their low pixel resolution and high cost. However, the second generation Microsoft Kinect™ (Kinect v2), an example of a consumer TOF camera (CTC), has superior technical characteristics at an affordable price. Since there is a patent on this device, the details of its functionality are not openly available; however, Lachat et al. (2015) assume that the technique behind this CTC is continuous wave modulation TOF.

2.2.3 Interferometry

Interferometry is the most accurate of the basic principles, with accuracies in the nanometre range. The basic operation of the interferometer consists of splitting a coherent light beam into two, one of which is projected towards a reference mirror while the other is projected towards a sample. Both rays are then reflected back to the beam splitter and projected towards a sensor for integration, where the phase shift between the beams is used to determine the relative depth (Figure 2-4). Optical Coherence Tomography (OCT) is an interferometric technique able to produce a full tomographic or topographic image depending on the penetration depth of the light. Interferometric techniques are classified into multiwavelength, holographic, speckle interferometry, and white light.

Reference surface

Measured surface

Beam splitter

Light source

$Z=\frac{n\lambda}{2}$

Image sensor

$Z=?$
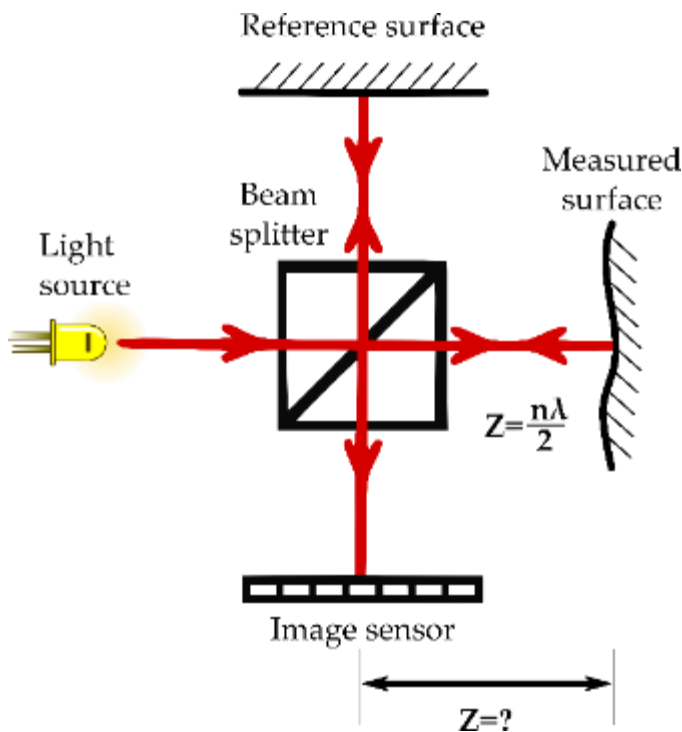
Figure 2-4:   Schematic representation of a Michelson interferometer where the relative depth "Z" is directly proportional to the wavelength of the light source "λ" and to the number of fringes "n".

2.2.4 Comparison of the most common 3-D vision techniques

Since this review paper deals with agriculture, it is important to analyse the most common implementations of the different 3-D vision techniques based on the basic principles. Table 2-2 presents the advantages and disadvantages of the most common sensor implementation described in the basic principles. The content of Table 2-2 was formed from an agricultural-based perspective. In the table it is clear that interferometric sensors are very scarce compared to

triangulation and TOF ones. Also, the technical characteristics of state-of-the-art sensors are improving the traditional disadvantages of the well-stablished commercial versions. This is particularly noticeable in the modern TOF cameras. CTCs have a good price/performance ratio, thus, they have a lot of potential in agriculture. Smart stereo vision sensors are increasingly common and are able to stream real-time depth measurements.

2.3 Applications in agriculture

2.3.1 Vehicle navigation

The agricultural sector is a pioneer in autonomous navigation relying on global navigation satellite system (GNSS). However, GNSS is not available in all agricultural environments at all times. Reactive sensor-based autonomous navigation is based on detailed information regarding the structures surrounding a machine such as objects, field structures, natural or artificial markers, and obstacles. In these type of applications, a superior perception is usually required, and 3-D imaging provides more information about the previously mentioned surrounding structures compared with 2-D. Automated and robotic systems could have faster acceptance by farmers if their safety aspects are well fulfilled (Griepentrog et al., 2009). Several reviews (Ji et al., 2009; Mousazadeh, 2013; Shalal et al., 2013) on autonomous navigation of agricultural vehicles have been written, however, there was little focus on the 3-D vision approach.

Basic Principle	Sensor/Technique	Advantages	Disadvantages
Triangulation	Consumer triangulation sensor (CTS)	-Off-the-shelf -Low cost -Provide RGB stream -Good community support, good documentation -Open source libraries available	-Vulnerable to sunlight, where no depth information is produced -Depth information is not possible at night or in very dark environments -Not weather resistant -Warm-up time required to stabilize the depth measurements (~1 h)
	Stereo vision	-Good community support, good documentation -Off-the-shelf smart cameras (with parallel computing) available -Robust enough for open field applications	-Low texture produce correspondence problems -Susceptible to direct sunlight -Computationally expensive -Depth range is highly dependent on the baseline distance
	Structure-from-motion	-Digital cameras are easily and economically available -Open source and commercial software for 3-D reconstruction -Suitable for aerial applications -Excellent portability	-Camera calibration and field references are a requirement for reliable measurements -Time consuming point cloud generation process is not suitable for real-time applications -Requires a lot of experience for obtaining good raw data
	Light sheet triangulation	-High precision -Fast image data acquisition and 3-D reconstruction -Limited working range due to the focus -Do not depend on external light sources -New versions have light filtering systems that allow them to handle sunlight	-High cost -Susceptible to sunlight -Time consuming data acquisition

		Advantages	Disadvantages
TOF	TOF camera	-Active illumination independent of an external lighting source -Able to acquire data at night or in dark/low light conditions -Commercial 3-D sensors in agriculture are based on the fast-improving photonic mixer device (PMD) technology -New versions have pixel resolutions of up to 4.2 Megapixels -New versions have depth measurement ranges of up to 25 m	-Most of them have low pixel resolution -Most of them are susceptible to direct sunlight -High cost
	Light sheet (pulse modulated) LI-DAR	-Emitted light beams and are robust against sunlight -Able to retrieve depth measurements at night or in dark environments -Robust against interference -Widely used in agricultural applications -Many research papers and information available -New versions perform well in adverse weather conditions (rain, snow, mist and dust)	-Poor performance in edge detection due the spacing between the light beams -Warm-up time required to stabilize the depth measurements (up to 2.5 h) -Normally bulky and with moving parts -Have problems under adverse weather conditions (rain, snow, mist and dust)
Interferometry	Optical coherent tomography (OCT)	-High accuracy -Near surface light penetration -High resolution	-High cost -Limited range -Highly-textured surfaces scatter the light beams -Relative measurements -Sensitive to vibrations -Difficult to implement

Table 2-2: Advantages and disadvantages of the most common sensor implementations, based on the basic principles for 3-D vision.

2.3.1.1 Triangulation

Autonomous navigation based on stereo vision was successfully achieved in several research studies using different cues. For example, in crop rows, cut-uncut edges, ridges, furrows, artificial markers, swaths and even stubble can be used. Kise et al. (2005) developed a stereo vision system that uses the 3-D crop row structure for automated guidance; problems like high computational load and blank pixels of some locations (particularly the ones that are further away) were reported, but were addressed by using a reduced resolution frame and filtering, respectively. Rovira-Más et al. (2007) used the cut-uncut edges of a maize field as a reference for autonomous guidance, they reported that a cloudy sky affected the 3-D point generation and the long maize leaves were blocking the camera, thus, recognizing the importance of the position of the camera. They also faced difficulties related with computational processing, but solved it by reducing the amount of points to a certain range.

(Hanawa et al., 2012) used ridges, furrows, and artificial markers for autonomous guidance emphasising the flexibility of their system. They reported limitations when the sunlight was very strong and the projected shadow of the tractor was in the range of the 3-D imaging system. (Blas and Blanke, 2011) developed an autonomous guidance system that uses a swath as the main reference. The 3-D imaging system had problems with the height resolution that failed to detect the difference on a very flat region of the swath. (Wang et al., 2011) relied on stereo vision to track the texture rich surface of a cultivated stubble field and calculate the vehicle's lateral offset. When the vehicle travelled in straight path, the maximum absolute deviation measured was 50 mm, and although it did not perform well in curved paths, no technical limitations regarding the stereo vision acquisition system were reported. Trinocular vision allows multiple baseline longitudes that complement each other for a more accurate depth measurement at different ranges. (Reina and Milella, 2012) used a trinocular vision and machine learning for ground or non-ground labelling in agricultural environments, reporting a classification precision of 91%. Furthermore, they fused stereo vision with other imaging techniques stating that the combination LIDAR-stereo vision is mutually complementary in many aspects. The classification results were better with the combined sensors than with the single sensors (Reina et al., 2015).

Tasks such as deformable and rigid obstacle recognition, reliability and operator protection are all considerable with 3-D sensors. Wei et al. (2005) developed and obstacle detection system using stereo vision to enhance the safety of autonomous vehicles stating the robustness against

foreground and background colours, and the limitation regarding the field of view and the number of tracking obstacles. Yang and Noguchi (2012) used two omnidirectional cameras to develop a human-detection safety system capable of acquiring a depth image with a reported error of less than half a meter, however, the experiment detects not more than a single human at day time.

(Nissimov et al., 2015) developed and obstacle detection system for greenhouse environment using a CTS with only few false positive detections and claiming that it could be used in a computer with limited processing capabilities. Several sources of error were mentioned such as problems with smooth and shiny surfaces, misalignment between the RGB and depth image, time delay (30 s) for a stable depth measurement after a quick rotation, synchronization, and mismatch between the RGB and depth images' field of view and point of view.

Night-time farming is being investigated in several publications since it provides a convenient environment for image acquisition, and potentially reducing the hazard of an autonomous vehicle colliding with humans. Another advantage is that fruit harvesting robots could become profitable if they are able to provide a 24/7 service. CTCs have a lot of potential for night farming applications since they are able to output an infrared stream. Kaizu and Choi (2012) developed an augmented reality 3-D system to assist tractor navigation at night-time. Although it was mainly developed using surveying and blending calibrated video images with computer graphics, it became clear that 3-D sensors could fit well to perform the same task.

2.3.1.2 TOF

Choi et al. (2014) developed a navigation system for a combine harvester based on a LIDAR (pulse modulation) mounted on a pan-tilt system that performs a 21° pitching. The system was evaluated under static and dynamic conditions with lateral root mean square (RMS) errors of 0.02 m and 0.07 m, respectively. Yin et al. (2013) used a TOF (continuous wave modulation) camera as the main navigation sensor for an agricultural vehicle that targets and follows a human, to further complement the concept of collaborative master-slave and multi-robot systems.

Commercial automatic guidance systems based on 3-D vision already exist in agriculture: CLAAS developed a smart 3-D stereo vision camera called CAM PILOT (CLAAS, 2014) that tracks different agricultural patterns such as ridges, swaths, crop rows and vineyards using 2-D and 3-D image processing techniques independently or in combination. Also, IFM electronic offers a smart continuous wave modulation TOF 3-D sensor (O3M system) where the emitter is situated in a separate unit from the receiver. It was specifically designed for outdoor use and

interferences such as sunlight or materials with different reflective characteristics do not influence the repeatability of the measured data. This systems is able to detect a swath's contour lines for automatic navigation, and it also provides automatic object recognition of 20 different objects in a range of up to 35 m (ifm electronics, 2016). Regarding driver assistance systems, CLAAS offers a stereo vision system for automatic trailer fill, called AUTO-FILL (CLAAS, 2009), for a forage harvester. The camera locates the trailer, tracks the crop jet and hit point, and determines the fill level. An equivalent system called IntelliFill (New Holland, 2010) was developed by New Holland (Turin, Italy) but uses instead a TOF camera. Hernandez et al. (2015) have experimented with a stereo pair from a different perspective using a UAV mounted with a camera system that was previously calibrated, and although the accuracy was not outstanding, the processing speed for updating the results was appealing, however, issues like the influence of wind in the stability of the UAV in the open field and thus the quality of the images (blur) still need to be evaluated. Finally, Naio Technologies (Ramonville-Saint-Agne, France) has developed a commercial field robot (Naio Technologies, 2014) for mechanical weeding that relies on stereo vision for autonomous navigation between the crop rows. Initially, they relied on a light sheet LIDAR (pulse modulation) for navigation within the crop rows, then upgraded the vision system to stereo vision, claiming that with it, they have a more accurate positioning and behaviour of the field robot and are able to detect smaller plants.

2.3.2   Crop husbandry

Important parameters like crop growth status, biomass estimation, height, shape, nutrient supply, and health status are better analysed using 3-D sensors since the acquired data can be used for measuring or correlating the previously mentioned crop parameters. If the data is geo-referenced, individual crop plant treatment can be applied.

Recently, Li et al. (2014) reviewed the state-of-the-art in plant phenotyping, where different 3-D vision techniques are used. It was concluded that their refinement and development will accelerate the phenotyping process. Rosell and Sanz (2012) reviewed the geometric characterization of tree crops, and Wulder et al. (2012) of forest trees where airborne LIDAR has become an important tool for characterization. Vos et al. (2007) reviewed plant modelling (virtual plants) which becomes increasingly important for conducting virtual experiments that otherwise would take years to perform in field conditions. This is closely related with the quantification of plant properties.

A number of reviews related with quality inspection and grading of agricultural products were conducted. Moreda et al. (2009) contributed with a review on the different vision technologies for size determination and grading. They expect 2-D imaging to be increasingly substituted by 3-D, and consider 3-D multispectral scanning (combination of multispectral data with 3-D surface reconstruction) a promising technology. Bac et al. (2014) reviewed harvesting robots for high-value crops where commercial versions are already available for strawberry harvesting, but the price/performance ratio still needs to improve in order to gain acceptance by farmers. A recently concluded research project ("CROPS," 2016) showed the recurrence in the use of 3-D sensors in harvesting robots.

2.3.2.1 Triangulation

Triangulation based techniques using UAVs have been thoroughly investigated. Since these aerial vehicles are very cost effective compared to airplanes, and the images are of higher resolution compared with satellites, more research and commercial applications can be foreseen. Structure from motion (SfM) has a big potential in aerial applications involving UAVs, which will be increasingly integrated in future agricultural practices, replacing solutions like satellite or manned aircraft. A number of open source software for 3-D reconstruction are available such as 123D, ARC3D, Photosynth, Visual SFM, Bundler + PMVS2 and MicMac. Recently, Jay et al. (2015) developed an automatic platform for open-field phenotyping using MicMac. They found that SfM is a convenient technique since intrinsic camera parameters are automatically estimated, therefore, camera calibration is not required. However, they encountered problems like occlusion and plant changing position from one image to the other due to the wind. Santos and Oliveira (2012) combined SfM with multi-view stereo to produce dense point clouds of a basil specimen (Ocimum bacilicum) and an ixora specimen (Ixora coccinea) for indoor plant phenotyping, however, they stated that the method is limited to not too dense plant canopies due to occlusion and matching problems, aside from the time-consuming image acquisition. The generation of DEMs by means of SfM is an increasingly common practice, and yet it remains unexplored, where useful information can be obtained for soil erosion, hydrological phenomena, and gullies monitoring (Martínez-Casasnovas et al., 2013). Zarco-Tejada et al. (2014) used and off-the-shelf colour camera, without the infrared filter, mounted on a UAV to acquire high resolution (5 cm·pixel^{-1}) DSMs for canopy height quantification. They obtained R2 determination coefficients of up to 0.8 compared with the manual measurements, proving that this inexpensive method can provide accuracies as good as the more costly airborne LIDAR. The potential disadvantage of the methodology is the high image overlapping and the low altitude requirements of UAVs flights. Geipel et al. (2014) acquired aerial images from a UAV and

combined 3-D shape information with the RGB spectral information for estimating corn grain yield. They obtained R2 determination coefficients of up to 0.74 using three different linear regression models, stating that dense point cloud generation requires high computational power, therefore, they downscaled the images by a factor of two.

Automated crop yield estimation, particularly in orchards and open field scenarios, is of great interest because is a very important parameter for farm management. It is a time consuming and labour intensive activity suitable for automation. Herrero-Huerta et al. (2015) proposed an automatic process for ground vineyard yield estimation by acquiring five images with an off-the-shelf camera and reconstructing the grape clusters (using SfM technique) at a close range, reporting that the main constrains depend on weather conditions and suggesting the use of artificial light and light diffusers to overcome them. Moonrinta et al. (2010) also proposed a method for pineapple plantations, considering SfM a promising technique, but recognizing that more work needs to be done to increase the accuracy of their recognition and tracking pipeline. Wang et al. (2012) used an autonomous stereo vision system for yield estimation of an apple orchard that works at night using ring flashes to illuminate the scene, they reported problems due to occlusion, specular reflections, colour heterogeneity, and a bias in the shape-from-stereo algorithm that caused the apple location to be estimated closer to the camera. To solve the bias problem, they placed artificial landmarks every three trees to recalibrate the vision system.

3-D models of trees, plant and agricultural structures are on demand as they help substitute difficult and expensive experiments. Model-based design reduces costs by avoiding redesign and removing the necessity to build a real prototype for experiment and evaluation (Arikapudi et al., 2015). Virtual simulation requires 3-D information of agricultural structures to create a model based on real information in order to pave the way for the following robotic application like harvesting, thinning, pruning, etc. Plant phenotyping is very important for plant breading, not just for increasing productivity, but also for minimising the effects of global warming in future farming. Table 2-3 shows that most of the autonomous phenotyping platforms are for research purposes and rely on 3-D vision (mainly triangulation). It can be seen that shadowing devices are commonly used to maintain constant lighting conditions. The ones that do not have, it is because they are in indoor or greenhouse environments where light can be easily controlled.

Platform	Basic Principle	Shadowing Device	Environment	Institution	Type
Becam (UMR Itap, 2015)	Triangulation	✓	Open field	UMR-ITAP	Research
BoniRob (Deepfield Robotics, 2016)	TOF	✓	Open field	Deepfield Robotics	Commercial
BredVision (Busemeyer et al., 2013)	TOF	✓	Open field	University of Applied Sciences Osnabrück	Research
Heliaphen (Optimalog, 2014)	Triangulation	✗	Greenhouse	Optimalog	Research
Ladybird (The University of Sidney, 2015)	TOF and Triangulation	✓	Open field	University of Sidney	Research
Marvin (Koenderink et al., 2009)	Triangulation	✓	Greenhouse	Wageningen University	Research
PhenoArch (INRA, 2014)	Triangulation	✓	Greenhouse	INRA-LEPSE (by LemnaTec)	Research
Phenobot (Polder et al., 2013)	TOF and Triangulation	✗	Greenhouse	Wageningen University	Research
PlantEye (Phenospex, 2013)	Triangulation	✗	Open field, Greenhouse	Phenospex	Commercial
Robot gardener (Alenyà et al., 2012)	Triangulation	✗	Indoor	GARNICS project	Research
SAS (Alci visionics & robotics, 2015)	Triangulation	✗	Greenhouse	Alci	Commercial
Scanalyzer (LemnaTec, 2015)	Triangulation	✗	Open field, Greenhouse	LemnaTec	Commercial
Spy-See (Polder et al., 2009)	TOF and Triangulation	✗	Greenhouse	Wageningen University	Research
Zea (BLUE RIVER TECHNOLOGY, 2015)	Triangulation	✓	Open field	Blue River	Commercial

Table 2-3: Autonomous platforms for reducing the time-consuming and repetitive phenotyping practice.

Shape-from-silhouette (SfS) has proved to be effective in characterizing the often complex plant architecture. Noordam et al. evaluated several 3-D image acquisition techniques and selected the reverse volumetric intersection technique, which is related to SfS, to obtain the best model of a rose with overlapping leaves for the 3-D vision system of a robotic rose cutter. They found this technique attractive because the addition of more cameras from different angles resulted in more information, and to a lesser extent in more processing (if multiple cameras are considered).

The task of rose cutting consists in locating the stem and trace it down until the cutting position by taking multiple images to generate the depth image (Hemming et al., 2005). (Tabb, 2013) tried to reconstruct trees based on the SfS technique taking into consideration that the trees do not contain concavities, but some noisy regions where present and post-processing filtering was required. (Billiot et al., 2013) evaluated Shape-from-focus technique using a monocular camera and two power LEDs. They also developed an acquisition platform that performs a controlled displacement perpendicular to the ground to acquire a stack of 3-D images of wheat ears. The focus value of each pixel in every 2-D image was used to obtain the depth information. They considered that the indoor system needs further development, but it can be translated to open field applications like crop characterization and yield estimation.

Jin and Tang (2009) developed an interplant spacing measuring system for corn plants at an early stage, using stereo vision. The system is able to detect almost 96% of the corn plants but with less accuracy at detecting the centre position (62% to 74%) with a processing time between 5 and 20 s. Zhao et al. (2012) used a light beam triangulation sensor mounted on a 2-D scanning platform to obtain the 3-D shape of zucchini (Cucurbita pepo) leaves, for detecting the water stress of the plant by tracing its morphological traits (wilting). They consider that little research has been done using a 3-D approach, even though it provides more reliable information about the wilting behaviour in response to water stress. Piron et al. (2011) used structured light volume sequentially coded technique to discriminate between weed and crop plants at an early stage by using the difference in height as the main cue.

They reported several problems such as limited projector depth of field, high dynamic range scene, internal reflections, thin objects and occlusions. Additionally, they also reported the solutions for every problem. Lino and Dal Fabbro (2001) used structured light volume shadow Moiré to reconstruct the 3-D shape of pears. The results were compared with the more precise

light sheet triangulation technique to evaluate its accuracy by correlating the depth measurements with R2 determination coefficient between 0.93 and 0.99, they reported visual noise in a small region but failed to explain the reason of it.

(Šeatović et al., 2009) used a light-sheet triangulation sensor (smart 3-D camera) for the development of a real-time broad-leafed weed (Rumex obtusifolius) detection and herbicide spraying prototype in grasslands. The detection rate was high, but decreased when clover or other broad-leafed plants were present. They concluded that a 3-D approach offers by far a more robust segmentation (with the help of height information) and classification of the leaves compared with a 2-D. (Wolff, 2012) developed an open field light-sheet triangulation system for plant phenotyping that consists of two cameras to reduce occlusion, arranged in an enclosed platform. The system scans around 2200 plants per working day, but requires two operators for moving the platform between each acquisition process. The combination of 3-D shape and spectral information are useful for farm management, and they can be acquired either from the ground (Kise and Zhang, 2008) or from the air (Rovira-Más et al., 2005). Strothmann et al. (2014) fused three light sheets (405, 532, and 650 nm) in a triangulation system using a single camera allowing not just 3-D shape reconstruction, but also to gain reflectance information.

Innovative applications have been developed with stereo vision like the inside tyre-terrain contact profile measurement of an off-road vehicle (Guthrie et al., 2014), which can be potentially useful for agricultural machinery testing and soil compaction analysis. The authors mention the complexity of the preparatory steps before depth calculation such as camera calibration, stereo rectification, correspondence problem, 3-D point computation and point cloud scaling. Shape-from-shading technique has been used to improve a common problem of 2-D vision systems for apple quality grading, where the stem-end or calyx could be incorrectly classified as a defect. Jiang et al. (2009) compared a traditional 2-D detection approach with a 3-D vision system based on Shape-from-shading. The result was a decrease of 30% in the overall error rate by using the 3-D approach, however, a zigzag effect at the apple's boundary was generated (in interlaced video) due to the high speed of the apples on the conveyor.

Advances in technology have allowed the creation of new devices inspired from old theoretical concepts, which is the case of the development of commercial light field cameras. They provide 4-D light field information, allowing 3-D reconstruction, and were evaluated by Möller et al. (2016) for cereal phenotyping in open field using a multi-sensor platform (Busemeyer et al., 2013). Polder and Hofstee (2014) also evaluated the same light field camera (Raytrix GmbH, Kiel, Germany) for tomato plant phenotyping in a greenhouse (see Figure 2-5), stating several

disadvantages such as computationally (depth image calculation) and economically expensive, one aperture setting, and limited field of view. Vision Robotics (2014) developed a robotic pruning prototype for grapevines based on stereo vision. The enclosed system was designed to control the lighting and protect the two robotic hands with cutter end-effectors, but no information is available regarding the technical limitations of the system.
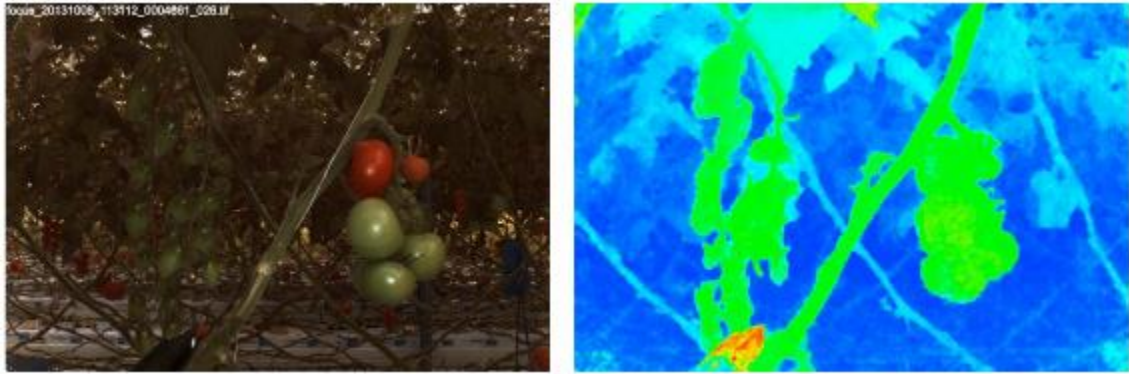
Figure 2-5:   RGB (left) and depth image (right) using a light field camera (reproduced from Polder and Hofstee (2014)).

Although not many effective technologies are available today, soil sensing is of high importance. The vertical dimension of soil properties is of high interest and hence, 2-D sensors can only provide information about structures at the field surface. For tillage operations, it is increasingly important to know about soil roughness and 3-D vision sensors provide a fast acquisition solution. Interestingly, under ideal light conditions Marinello et al. (2013) successfully measured the soil roughness using a CTS, under favourable lightning conditions, highlighting the effects of oversaturation (due to excessive sunlight exposure) as the main limitation.

2.3.2.2 TOF

With regard to the use of scout robots, Garrido et al. (2015) reconstructed maize plants by overlapping LIDAR point clouds using a field robot for data acquisition and a robotic total station for geo-referencing the point clouds, relying on sensor fusion, filtering and processing to reconstruct the 3-D plant structure, and concluding that the orientation of the 3-D sensor is very important.

Weiss et al. (2010) used a low pixel resolution 3-D LIDAR (pulse modulation) to evaluate different machine learning classifiers in an indoor environment. They achieved a classification precision of nearly 99% with one of the trained classifiers (simple logistic regression) using plants of six different species. Afterwards, they used the same 3-D LIDAR for plant detection

and mapping in outdoor conditions with a plant detection rate of about 60%. Even though the plant detection rate in the open field was not outstanding, and the sensor's pixel resolution was poor (29×59 pixels), the authors emphasised the advantages of the sensor (reliability under different light and weather conditions) and considered it as the most promising sensor technology for agricultural robotics (Weiss and Biber, 2011).

Saeys et al. (2009) evaluated two light-sheet LIDARs, a continuous wave and a pulse modulated, to estimate wheat ear density and crop volume mounted on a combine harvester.

They successfully predicted the crop density by conducting experiments with different crop densities (controlled), speeds and vibrations. The LIDAR hits were used to reconstruct the 3-D field in post processing but better results were obtained using the continuous wave LIDAR since its scanning rate were intrinsically higher than the pulse LIDAR.

Nakarmi and Tang (2012) developed a system for sensing inter-plant spacing, using a state-of-the-art TOF camera (continuous wave modulation), that was fully covered to protect it from direct sunlight and wind. They mention the superiority of TOF cameras compared to a conventional stereo vision sensor, but also their limitations like the small field of view and the low pixel resolution. Adhikari and Karkee (2011) developed a 3-D imaging system for automatic pruning to identify unwanted branches and locate pruning points in a trained apple orchard. 90% of the pruning points were correctly located using a TOF camera (continuous wave modulation).

Gongal et al. (2014) investigated the fusion of a 2-D camera with a TOF camera (continuous wave modulation) for apple yield estimation in trained orchards. They were able to recognize 88% of the apples emphasizing the significant increase of visibility when the images are captured from both sides of the tree canopies, rather than just from one side. But they acknowledged that the major challenge was the limited visibility of apples, where some of them were completely occluded by leaves and branches. Tanaka et al. (2014) used a light-beam triangulation scanner to acquire the 3-D shape of a rotary tiller blade to investigate if a 3-D printed replica made of resin could perform similarly for low-cost prototyping in the effort of improving the design/performance of the rotary tiller blade.

Vázquez-Arellano et al. (2016) have emphasised the flexibility and possibilities of CTCs in agricultural applications for different agricultural environments and conditions, since they were able to reconstruct and geo-reference maize plants in a greenhouse and open field environments under different lighting conditions (Figure 2-6). Deepfield Robotics is also relying on a CTC

for plant phenotyping in their Bonirob field robot (Deepfield Robotics, 2016), but for acquiring stable images, they rely on a shadowing device that houses the CTC together with an artificial light source.
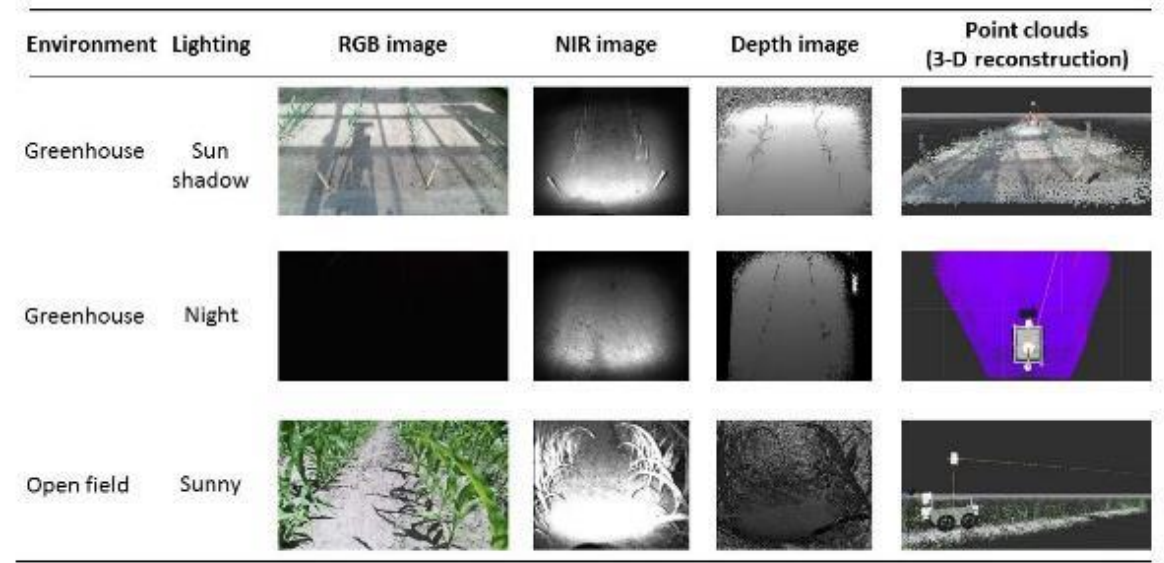
Figure 2-6:   Reconstruction of maize plants using a CTC mounted on a field robot in different agricultural environments (reproduced from Vázquez-Arellano et al. (2016)).

2.3.2.3 Interferometry

Interferometry has been used for a long time to measure plant growth or motion changes under different stimuli. Currently, interferometric techniques are being investigated using a 3-D approach for seed inspection and quality control. There is a worldwide rush for preserving genetic pool of different crops in seed banks for future breeding requirements. Lee et al. (2011) used optical coherence tomography (OCT) based on white-light interferometry to detect infected melon seeds, until now, it is one of the few examples of 3-D reconstruction using interferometry in agriculture (Figure 2-7).
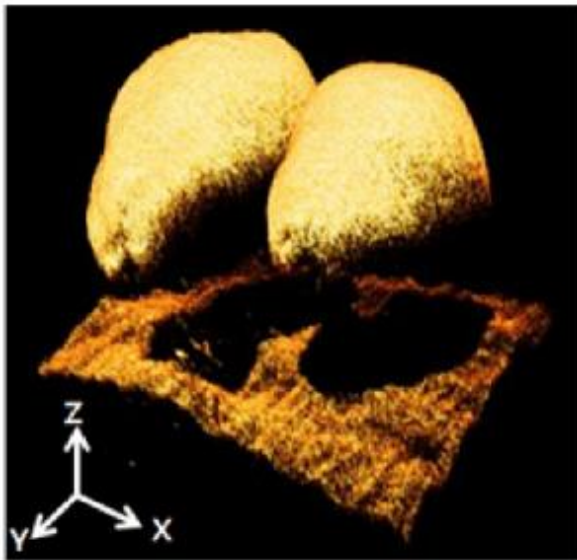
Figure 2-7:   3-D reconstruction of melon seeds based on interferometry (reproduced from Lee et al. (2011)).

Later on, they did the same for cucumber seeds (Lee et al. 2012). Barbosa and Lino (2007) used electronic speckle interferometry (ESPI) for 3-D shape measurement of a peach concluding that the technique is promising for quality control of agricultural products with smooth and delicate tissue. Madjarova et al. (2003) also used ESPI for flower blooming growth analysis that can provide useful information of the effects changing weather patterns in flowering (which is sensitive to temperature variability) and thus crop production. This study shows that a high resolution camera is important to resolve high fringe densities. Plant movement analysis is other application area, where Fox and Puffer (1977) relied on holographic interferometry to measure motion changes undergone by a mature Stapelia variegate under phototropic stimuli, where a reference object was used to detect unwanted movements. In the other hand, Thilakarathne et al. (2014) relied on white-light interferometry to measure the nanometric intrinsic fluctuations of rice (Oriza sativa L.) exposed to different ozone concentrations to investigate the damage and recovery of the plant. They explained that the usage of interferometry was limited because of two main factors: complexity of implementation and the optical properties of the plant itself (highly scattering surface).

2.3.3 Animal husbandry

2.3.3.1 Triangulation

Milking robots have been in operation for the last two decades maintaining a steady growth and until recently, they had been the only commercial robots available in agriculture. Current milking robots use light sheet triangulation to estimate the position of the teats, but some experiments have been evaluating alternatives. Ben Azouz et al. (2015) evaluated a stereo vision system together with thermal infrared acknowledging the difficulty of obtaining a robust disparity estimation particularly in areas of homogeneous colour or occlusion. Similarly, Akhloufi (2014) compared two TOF cameras and a CTS, with the latter giving the overall best results mainly because of its superior pixel resolution and colour output.

Early attempts to reconstruct a pig's surface using light-volume triangulation were done by Van der Stuyft et al. (1991) without achieving good resolution. They reported the potential of the technique, but also the necessity to have expert knowledge of vision algorithm development and hardware implementation. Ju et al. (2004) used high resolution cameras (4500×3000 pixel) to achieve pig surface resolutions of approximately 0.4 mm pixel^{-1} based on multi-view stereo, but difficulties arose with regard to the discrimination between foreground (pig) and background surfaces, as well as residual corrupted range measurements due to pig's background occlusion. Automation in slaughtering facilities has allowed a surge in the use of structured light techniques for meat cutting, grading, sorting, and yield calculation since this activity is mainly done in indoor environments where illumination conditions are controlled (Hinz, 2012).

Cattle monitoring was discussed by Viazzi et al. (2014), who compared a 2-D camera with a CTS for lameness detection concluding that the 3-D approach overcomes the limitations of the 2-D, however, CTSs have also limitations such as their sensitivity to sunlight and small field of view. Kawasue et al. (2013) used three CTS for evaluating the quality of cattle with accuracies of up to 93% compared with manual measurement. An important source of error was caused by the body hair of the cattle. Similarly, Kuzuhara et al. (2015) used a CTS to estimate parameters like biomass and milk yield of Holstein cows. They obtained R2 coefficient of determination of 0.8 for body weight and 0.62 for milk yield. They mention limitations regarding the sensitivity of the sensor to natural light, but this problem was solved by performing the experiments inside the cowshed.

Animal welfare, health monitoring, indoor navigation of robots, feeding and cleaning robots are some examples of the many tasks where 3-D sensors could be included in animal husbandry.

Several 3-D imaging techniques can be used since most of the time the measurements can be done in an indoor environment. Animal husbandry has experienced recent advances in quality evaluation and monitoring. There is a need to develop integrated monitoring systems that can measure important performance parameters including physiological variables like shape, size and weight. Menesatti et al. (2014) developed a low-cost stereo pair system for measuring such parameters for sheep with webcams, reporting R^2 determination coefficient of 0.8. Pallottino et al. (2015) used a similar system for measuring body traits for breeding Lipizzan horses, reporting a high correlation coefficient ($r = 0.998$) between manual measurements and stereo vision. Wu et al. (2004) developed a system for swine monitoring, using three high-resolution stereo vision systems (side, rear and top) for further integration and overall 3-D shape reconstruction. Aquaculture production currently supplies nearly 50% of the fish consumed in the world and the farmers in this sector face major difficulties in monitoring their fish stock. According to a review by Zion (2012), several stereo vision systems have been used for measuring individual fish dimensions and mass. Their further improvement could lead to the development of a system capable of estimating the overall biomass and monitoring of fish welfare. However, problems like fish occlusion and poor water transparency need to be addressed.

2.3.3.2 TOF

Research has been conducted on the development of a robot for herding milking cows (Underwood et al., 2013). It uses a 360° 3-D LIDAR as its main navigation sensor. Several aspects like response of the cows to the presence of a robot, remote controlled operation, and software algorithms for detecting and tracking were assessed. After three herding tests, the authors showed that remote herding was possible with the potential to improve animal welfare.

2.3.4 Summary

Table 2-4 lists the technical difficulties so far encountered in the reviewed papers, with the techniques based on the basic principles in agricultural applications. The table shows many comments on technical difficulties regarding stereo vision, and that is expected since it is economically affordable and has been fully studied compared with other techniques. Also, a common technical difficulty is the sensitivity against natural light, where shadowing devices are commonly used to maintain the illumination condition as constant as possible. Occlusion is also a very recurrent problem, here, 3-D sensor pose and position during data acquisition play an important role to minimize occlusions.

2.4 Discussion

New commercial 3-D capable optical sensors such as light field and polarization cameras have recently appeared. The latter provide information about the polarization properties of the image. This information is used by the Shape-from-polarization technique for 3-D reconstruction.

The previously mentioned sensors rely on triangulation techniques that were not included in the classification proposed by Schwarte et al. (1999). The Austrian Institute of Technology (Belbachir et al., 2014) has recently developed a dynamic stereo vision camera that continuously rotates to generate a real-time 360° 3-D view. This camera exploits the high sampling rate and low latency capabilities of the dynamic vision sensor (DVS) that only senses changes at a pixel-level, caused by movement, significantly reducing the amount of acquired data. Parallel computation, accelerated by the incorporation of field programmable gate arrays (FPGAs), has enabled the emergence of 3-D smart cameras that have embedded processing. In the case of triangulation based sensors, the depth map can be generated as an output stream, thus, real-time measurements can be performed. Velodyne (2015) has released a 360° 3-D LIDAR that is relatively cost-effective and compact in size. Since this type of sensor is preferred in military and automotive autonomous applications, its presence could be also expected in future agricultural scenarios.

Several trends have been detected during this extensive literature review. Any 3-D vision sensor has its disadvantages and advantages, therefore, complementary 3-D sensor fusion provides a more robust performance depending on the application. Also, there is no 3-D sensor completely immune to noise sources such as natural light, sunlight intensity variations, adverse weather conditions (rain, snow, mist and dust), and light reflectivity differences (due to colour and texture). The outdoor range and resolution of TOF cameras are both expected to increase, while the cost per pixel decreases (Frey, 2010). Several TOF cameras have been tested and compared for agricultural applications by Kazmi et al. (2014) and Klose et al. (2011), and although there are still several noise sources that affect them, the technical characteristics of some versions are outstanding and efforts to extend the range of TOF cameras are taking place. Higher resolution TOF cameras (pulse modulation) are already commercially available and reaching resolutions of 4.2 Megapixels (odos imaging, 2014). Also, higher measurement ranges are commercially available like the previously mentioned smart TOF 3-D sensor (continuous wave modulation) by IFM electronics that has a measuring range of up to 35 m. Consumer 3-D sensors have had a big impact in agricultural automation and robotics research even with their outdoor limitations, where shading devices have so far minimized the problem.

Basic Principle	Technique	Application	Technical Difficulties
Triangulation	Stereo vision	-Autonomous navigation (Hanawa et al., 2012; Kise et al., 2005; Rovira-Más et al., 2007), (Wang et al., 2011), (Reina et al., 2015), (Yang and Noguchi, 2012) -Crop husbandry (Wang et al., 2013), (Rovira-Más et al., 2005), (Guthrie et al., 2014) -Animal husbandry (Ben Azouz et al., 2015), (Zion, 2012)	-Blank pixels of some locations specially the ones that are further away from the camera -Low light (cloudy sky) affects 3-D point generation -Direct sunlight and shadows in a sunny day affect strongly the depth image generation -Uniform texture of long leaves affect the 3-D point generation -Limited field of view -External illumination is required for night implementations -Correspondence and parallax problems -A robust disparity estimation is difficult in areas of homogeneous colour or occlusion -Specular reflections -Colour heterogeneity of the target object -A constant altitude needs to be maintained if a stereo vision system is mounted on a UAV -Camera calibration is necessary -Occlusion of leaves -Selection of a suitable camera position
	Multi-view stereo	-Crop husbandry (Santos and De Oliveira, 2012) -Animal husbandry (Ju et al., 2004)	-Surface integration from multiple views is the main obstacle -Challenging software engineering if high-resolution surface reconstruction is desired -Software obstacles associated with handling large images during system calibration and stereo matching
	Multiple-baseline stereo	-Autonomous navigation (Reina and Milella, 2012)	-Handling a rich 3-D data is computationally demanding

	Structure-from-motion	-Crop husbandry (Jay et al., 2015), (Geipel et al., 2014; Herrero-Huerta et al., 2015; Zarco-Tejada et al., 2014), (Kise and Zhang, 2008)	-Occlusion of leaves
			    -Plant changing position from one image to the other due to the wind
			    -High computation power is required to generate a dense point cloud
			    -Determination of a suitable Image overlapping percentage
			    -Greater hectare coverage requires higher altitudes when using UAVs
			    -The camera's pixel resolution determines the field spatial resolution
			    -Image mosaicking is technically difficult from UAVs due to the translational and rotational movements of the camera

	Shape-from-Silhouette	-Crop husbandry (Hemming et al., 2005; Noordam et al., 2005; Tabb, 2013)	-3-D reconstruction results strongly depend on good image pre-processing
			    -Camera calibration is important if several cameras are used
			    -Dense and random canopy branching is more difficult to reconstruct
			    -Post-processing filtering may be required to remove noisy regions

	Structured light (light volume) sequentially coded	-Crop husbandry (Piron et al., 2011)	-Limited projector depth of field
			    -High dynamic range scene
			    -Internal reflections
			    -Thin objects
			    -Occlusions

Triangulation	Structured light (light volume) pseudo random pattern	-Autonomous navigation (Nissimov et al., 2015)
		-Animal husbandry (Akhloufi, 2014), (Kuzuhara et al., 2015)
			-Strong sensitivity to natural light
			    -Small field of view
			    -Smooth and shiny surfaces do not produce reliable depth measurements
			    -Misalignment between the RGB and depth image due to the difference in pixel resolution
			    -Time delay (30 s) for a stable depth measurement after a quick rotation
			    -Mismatch between the RGB and depth images' field of view and point of view

	Shape-from-Shading	-Crop husbandry (Jiang et al., 2009)	-A zigzag effect at the target object's boundary is generated (in interlaced video) if it moves at high speeds

	Structured light shadow Moiré	-Crop husbandry (Lino and Dal Fabbro, 2001)	-Sensitive to disturbances (*e.g.*, surface reflectivity) that become a source of noise
	Shape-from-focus	-Crop husbandry (Billiot et al., 2013)	-Limited depth of field decreases the accuracy of the 3-D reconstruction
TOF	Pulse modulation (light sheet)	-Autonomous navigation (Choi et al., 2014) -Crop husbandry (Garrido et al., 2015), (Saeys et al., 2009)	-Limited perception of the surrounding structures -Requires movement to obtain 3-D data -Pitching, rolling or jawing using servo motors (*i.e.*, pan-tilt unit) is a method to extend the field of view, but adds technical difficulties -Point cloud registration requires sensor fusion -Small plants are difficult to detect -Lower sampling rate and accuracy compared to continuous wave modulation TOF
	Pulse modulation (light volume)	-Autonomous navigation and crop husbandry (Weiss et al., 2010)	-Limited pixel resolution -Difficulty to distinguish small structures with complex shapes
	Continuous wave modulation (light sheet)	-Crop husbandry (Saeys et al., 2009)	-Poor distance range measurement (up to 3m)
	Continuous wave modulation (light volume)	-Crop husbandry (Adhikari and Karkee, 2011; Gongal et al., 2014; Nakarmi and Tang, 2012) -Animal husbandry (Akhloufi, 2014)	-Small field of view -Low pixel resolution -Calibration could be required to correct radial distortion -Requires a sunlight cover for better results -Limited visibility due to occlusion -Lack of colour output that could be useful for a better image segmentation
	White-light	-Crop husbandry (Lee et al. 2011; Lee et al. 2012),	-The scattering surface of the plant forms speckles that affect the accuracy -Complexity of implementation

Inter-fer-ometry			
		(Thilakarathne et al., 2014)	
	Holographic	-Crop husbandry (Fox and Puffer, 1977)	-Need of a reference object in the image to detect disturbances
	Speckle	-Crop husbandry (Barbosa and Lino, 2007), (Madjarova et al., 2003)	-Agricultural products with rough surface could be difficult to reconstruct -High camera resolutions provide better capabilities to resolve high fringe densities

Table 2-4: Summary of the technical difficulties of the 3-D techniques used in agricultural applications.

2.5 Conclusions

Currently, 3-D sensors are becoming smaller, smarter, and cheaper. Therefore, technology breakthroughs are already possible if enough research were commercialized, a statement justified by the fact that some commercial implementations in agriculture are mentioned in this paper. Since agricultural environments are substantially complex, 3-D vision can play an increasing role in enhanced perception that could be suitable in a number of applications in every agricultural scenarios. The true value of 3-D data lies precisely in the superior sensing capabilities compared to 2-D. Several market forecasts have recently appeared around topics such as LED lighting, 3-D sensors, UAVs, and agricultural automation and robotics. Almost all of them forecast a profitable market for these interconnected topics by the end of the decade. Some 3-D techniques have either not been tested (i.e., photometric stereo, shape-from-polarization, shape-from-zooming, flash LIDAR) or more research has to be done with the rich variety of 3-D imaging techniques in agricultural applications (i.e., interferometry, light field, CTCs). 2-D and 3-D fusion is very promising, since it takes the advantages of the two, and has proved to be useful for either obtaining more information about the object's surface or facilitating the image segmentation process. All 3-D sensors are sensitive in one way or another to sunlight, however, more research needs to be done to reduce its effects and stop relying on shading devices. As a matter of fact, a positive effect of this disadvantage is that autonomous night farming could be investigated more thoroughly since 3-D sensors behave properly in this environment.

This review presents the rich variety of 3-D imaging techniques that have not been tested, the potential of the ones that have already been tested (due to the increase of their technical characteristics), the working principle behind the 3-D imaging sensors, and the potential of consumer 3-D sensors in agricultural applications. The pace of change in 3-D imaging technology is accelerating, therefore, the possibilities of this technology are immense. Reasons like reduced labour availability, scarcity of natural resources, and consumer demand for quality products are driving the need for automation in agriculture. Since 3-D vision is a key technology for automation, more implementations are yet to come.

Author Contributions: The first author conducted the literature review with the discussion, correction, and guidance provided by the second, third, and fourth authors at every stage of the process: from the general structure to the specific details.

Conflicts of Interest: The authors declare no conflict of interest.

References

Adhikari, B., Karkee, M., 2011. 3D Reconstruction of apple trees for mechanical pruning. ASABE Annu. Int. Meet. 7004.

Akhloufi, M., 2014. 3D vision system for intelligent milking robot automation, in: Intelligent Robots and Computer Vision XXXI: Algorithms and Techniques. San Francisco, California, USA.

Alci visionics & robotics, 2015. Sampling Automation System: SAS [WWW Document]. URL http://www.alci.fr/ (accessed 2.1.16).

Alenyà, G., Dellen, B., Foix, S., Torras, C., 2012. Robotic leaf probing via segmentation of range data into surface patches, in: IROS Workshop on Agricultural Robotics: Enabling Safe, Efficient, Affordable Robots for Food Production. IEEE/RSJ, Vilamoura, Portugal, pp. 1–6.

Arikapudi, R., Vougioukas, S., Saracoglu, T., 2015. Orchard tree digitization for stuctural-geometrical modeling, in: Stafford, J. V. (Ed.), Precision Agriculture '15. Wageningen Academic Publishers, Wageningen, The Netherlands, pp. 329–336.

Bac, C.W., van Henten, E.J., Edan, Y., 2014. Harvesting Robots for High-value Crops : State-of-the-art Review and Challenges Ahead. J. F. Robot. 00, 1–24. doi:10.1002/rob

Barbosa, E.A., Lino, A.C.L., 2007. Multiwavelength electronic speckle pattern interferometry for surface shape measurement. Appl. Opt. 46, 2624–31.

Belbachir, A.N., Ieee, M., Schraml, S., Mayerhofer, M., Hofstätter, M., 2014. A Novel HDR Depth Camera for Real-time 3D 360 ° Panoramic Vision, in: IEEE Conference on Computer Vision and Pattern Recognition. IEEE, Columbus, USA, pp. 425–432.

Bellmann, A., Hellwich, O., Rodehorst, V., Yilmaz, U., 2007. A Benchmarking Dataset for Performance Evaluation of Automatic Surface Reconstruction Algorithms, in: IEEE Conference on Computer Vision and Pattern Recognition. IEEE, Minneapolis, USA, pp. 1–8. doi:10.1109/CVPR.2007.383349

Ben Azouz, A., Esmonde, H., Corcoran, B., O'Callaghan, E., 2015. Development of a teat sensing system for robotic milking by combining thermal imaging and stereovision technique. Comput. Electron. Agric. 110, 162–170. doi:10.1016/j.compag.2014.11.004

Bergerman, M., Henten, E. Van, Billingsley, J., Reid, J., Mingcong, D., 2013. IEEE Robotics and Automation Society Technical Committee on Agricultural Robotics and Automation. IEEE Robot. Autom. Mag. 20, 20–24. doi:10.1109/MRA.2013.2255513

Besl, P.J., 1988. Active , Optical Range Imaging Sensors. Mach. Vis. Appl. 01, 127–152.

Billiot, B., Cointault, F., Journaux, L., Simon, J.-C., Gouton, P., 2013. 3D Image Acquisition System Based on Shape from Focus Technique. Sensors 13, 5040–5053. doi:10.3390/s130405040

Blais, F., 2004. Review of 20 years of range sensor development. J. Electron. Imaging 13, 231. doi:10.1117/1.1631921

Blas, M.R., Blanke, M., 2011. Stereo vision with texture learning for fault-tolerant automatic baling. Comput. Electron. Agric. 75, 159–168. doi:10.1016/j.compag.2010.10.012

BLUE RIVER TECHNOLOGY, 2015. Zea [WWW Document]. URL http://www.bluerivert.com/ (accessed 2.1.16).

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. doi:10.3390/s130302830

Büttgen, B., Oggier, T., Lehmann, M., 2005. CCD/CMOS lock-in pixel for range imaging: challenges, limitations and state-of-the-art, in: 1st Range Imaging Research Day. Hilmar Ingensand and Timo Kahlmann, Zurich, Switzerland, pp. 21–32.

Cheung, K.G., Baker, S., Kanade, T., 2005. Shape-From-Silhouette Across Time Part I : Theory and Algorithms. Int. J. Comput. Vis. 62, 221–247.

Choi, J., Yin, X., Yang, L., Noguchi, N., 2014. Development of a laser scanner-based navigation system for a combine harvester. Eng. Agric. Environ. Food 7, 7–13. doi:10.1016/j.eaef.2013.12.002

CLAAS, 2014. CAM PILOT [WWW Document]. URL http://www.claas.de/produkte/easy/lenksysteme/optische-lenksysteme/cam-pilot (accessed 1.26.16).

CLAAS, 2009. AUTO FILL [WWW Document]. URL http://www.claas.de/produkte/easy/cemos/cemos-automatic (accessed 1.26.16).

CROPS [WWW Document], 2016. . Intell. Sens. Manip. Sustain. Prod. Harvest. high value Crop. clever Robot. Crop. URL http://cordis.europa.eu/result/rcn/90611_en.html (accessed 2.1.16).

Deepfield Robotics, 2016. BoniRob [WWW Document]. URL http://www.deepfield-robotics.com/ (accessed 2.1.16).

Edan, Y., Han, S., Kondo, N., 2009. Automation in agriculture, in: Nof, S.Y. (Ed.), Handbook of Automation. Springer, West Lafalette, USA, pp. 1095–1128.

Eddershaw, T., 2014. Imaging in depth. IMAGING Mach. Vis. Eur. 34–36.

Favaro, P., Soatto, S., 2005. A Geometric Approach to Shape from Defocus. IEEE Trans. Pattern Anal. Mach. Intell. 27, 1–12.

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. doi:10.1145/358669.358692

Fox, M.D., Puffer, L.G., 1977. Holographic interferometric measurement of motions in mature plants. Plant Physiol. 60, 30–3.

Frey, V., 2010. PMD Cameras for Automotive & Outdoor Applications. ifm electronic, Munich, Germany.

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. doi:10.3390/rs71215870

Geipel, J., Link, J., Claupein, W., 2014. Combined spectral and spatial modeling of corn yield based on aerial images and crop surface models acquired with an unmanned aircraft system. Remote Sens. 6, 10335–10355. doi:10.3390/rs61110335

Gongal, A., Amatya, S., Karkee, M., 2014. Identification of repetitive apples for improved crop-load estimation with dual-side imaging, in: ASABE and CSBE/SCGAB Annual International Meeting. Montreal, Canada, pp. 1–9.

Griepentrog, H.W., Andersen, N.A., Andersen, J.C., Blanke, M., Heinemann, O., Nielsen, J., Pedersen, S.M., Madsen, T.E., Wulfsohn, D., 2009. Safe and Reliable - Further Development of a Field Robot, in: 7th European Conference on Precision Agriculture (ECPA). Wageningen Academic Publishers, Wageningen, The Netherlands, pp. 857–866.

Grift, T., 2008. A review of automation and robotics for the bio- industry. J. Biomechatronics Eng. 1, 37–54.

Guthrie, A.G., Botha, T.R., Els, P.S., 2014. 3D computer vision contact patch measurements inside off-road vehicles tyres, in: 18th International Conference of the ISTVS. Seul, Korea, pp. 1–6.

Hanawa, K., Yamashita, T., Matsuo, Y., Hamada, Y., 2012. Development of a stereo vision system to assist the operation of agricultural tractors. Japan Agric. Res. Q. JARQ 46, 287–293. doi:10.6090/jarq.46.287

Hemming, J., Golbach, F., Noordam, J.C., 2005. Reverse Volumetric Intersection (RVI), a method to generate 3D images of plants using multiple views, in: Bornimer Agrartechn.

Berichte. Inst. für Agrartechnik e.V. Bornim, Potsdam-Bornim, Germany, pp. 17–27.

Hernandez, A., Murcia, H., Copot, C., De Keyser, R., 2015. Towards the Development of a Smart Flying Sensor: Illustration in the Field of Precision Agriculture. Sensors 16688–16709. doi:10.3390/s150716688

Herrero-Huerta, M., González-Aguilera, D., Rodriguez-Gonzalvez, P., Hernández-López, D., 2015. Vineyard yield estimation by automatic 3D bunch modelling in field conditions. Comput. Electron. Agric. 110, 17–26. doi:10.1016/j.compag.2014.10.003

Hinz, A., 2012. Objective grading and video image technology. E+V Technology GmbH & Co. KG, Oranienburg, Germany.

Horn, B.K.P., 1970. Shape From Shading: A Method for Obtaining the Shape of a Smooth Opaque Object From One View, Technical Report. Massachusetts Institute of Technology, Cambridge, USA.

ifm electronic, 2016. 3D Smart Sensor - your assistant on mobile machines [WWW Document]. URL http://www.ifm.com (accessed 1.22.16).

INRA, 2014. PhenoArch [WWW Document]. URL http://www.inra.fr/ (accessed 2.1.16).

Jähne, B., Haußecker, H., Geißler, P., 1999. Handbook of computer vision and applications. Academic Press.

Jarvis, R., 1983. A perspective on range finding techniques for computer vision. IEEE Trans. Pattern Anal. Mach. Intell. PAMI-5, 122–139. doi:10.1109/TPAMI.1983.4767365

Jay, S., Rabatel, G., Hadoux, X., Moura, D., Gorretta, N., 2015. In-field crop row phenotyping from 3D modeling performed using Structure from Motion. Comput. Electron. Agric. 110, 70–77. doi:10.1016/j.compag.2014.09.021

Ji, B., Zhu, W., Liu, B., Ma, C., Li, X., 2009. Review of recent machine-vision technologies in agriculture, in: Second International Symposium on Knowledge Acquisition and Modeling. Ieee, pp. 330–334. doi:10.1109/KAM.2009.231

Jiang, L., Zhu, B., Cheng, X., Luo, Y., Tao, Y., 2009. 3D surface reconstruction and analysis in automated apple stem-end/calyx identification. Trans. ASABE 52, 1775–1784.

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

Joergensen, R.N., 2002. Study on Line Imaging Spectroscopy as a Tool for Nitrogen Diagnostics in Precision Farming. The Royal Veterinary and Agricultural University, Copenhagen.

Ju, X., Siebert, J.P., McFarlane, N.J.B., Wu, J., Tillett, R.D., Schofield, C.P., 2004. A stereo imaging system for the metric 3D recovery of porcine surface anatomy. Sens. Rev. 24, 298–307. doi:10.1108/02602280410545948

Kaizu, Y., Choi, J., 2012. Development of a tractor navigation system using augmented reality. Engineeiting Agric. Environ. Food 5, 96–101.

Kawasue, K., Ikeda, T., Tokunaga, T., Harada, H., 2013. Three-Dimensional Shape Measurement System for Black Cattle Using KINECT Sensor. Int. J. Circuits, Syst. Signal Process. 7, 222–230.

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. doi:10.1016/j.isprsjprs.2013.11.012

Kise, M., Zhang, Q., 2008. Creating a panoramic field image using multi-spectral stereovision system. Comput. Electron. Agric. 60, 67–75. doi:10.1016/j.compag.2007.07.002

Kise, M., Zhang, Q., Rovira Más, F., 2005. A Stereovision-based Crop Row Detection Method for Tractor-automated Guidance. Biosyst. Eng. 90, 357–367. doi:10.1016/j.biosystemseng.2004.12.008

Klose, R., Penlington, J., Ruckelshausen, A., 2011. Usability of 3D time-of-flight cameras for automatic plant phenotyping. Bornimer Agrartech. Berichte 69, 93–105.

Koenderink, N.J.J.P., Wigham, M., Golbach, F., Otten, G., Gerlich, R., Zedde, H.J. Van De, 2009. MARVIN: high speed 3D imaging for seedling classification, in: European Conference on Precision Agriculture. pp. 279–286.

Kutulakos, K.N., Seitz, S.M., 2000. A theory of shape by space carving. Int. J. Comput. Vis. 38, 199–218. doi:10.1023/A:1008191222954

Kuzuhara, Y., Kawamura, K., Yoshitoshi, R., Tamaki, T., Sugai, S., Ikegami, M., Kurokawa, Y., Obitsu, T., Okita, M., Sugino, T., Yasuda, T., 2015. A preliminarily study for predicting body weight and milk properties in lactating Holstein cows using a three-dimensional camera system. Comput. Electron. Agric. 111, 186–193. doi:10.1016/j.compag.2014.12.020

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:10.5194/isprsarchives-XL-5-W4-93-2015

Lange, R., 2000. Time-of-flight distance measurement with solid-state image sensors in CMOS / CCD-technology. University of Siegen.

Lavest, J., Rives, G., Dhome, M., 1993. Three-dimensional reconstruction by zooming. IEEE Trans. Robot. Autom. 9, 196–207.

Lee, C., Lee, S.-Y., Kim, J.-Y., Jung, H.-Y., Kim, J., 2011. Optical sensing method for screening disease in melon seeds by using optical coherence tomography. Sensors (Basel).

11, 9467–77. doi:10.3390/s111009467

Lee, S.-Y., Lee, C., Kim, J., Jung, H.-Y., 2012. Application of optical coherence tomography to detect Cucumber green mottle mosaic virus (CGMMV) infected cucumber seed. Hortic. Environ. Biotechnol. 53, 428–433. doi:10.1007/s13580-012-0071-x

LemnaTec, 2015. Scanalyzer [WWW Document]. URL http://www.lemnatec.com/ (accessed 2.1.16).

Li, L., Zhang, Q., Huang, D., 2014. A review of imaging techniques for plant phenotyping. Sensors 14, 20078–20111. doi:10.3390/s141120078

Lino, A.C.L., Dal Fabbro, I.M., 2001. Fruit profilometry based on shadow Moiré techniques. Ciênc Agrotec 28, 119–125.

Lobay, A., Forsyth, D. a., 2006. Shape from Texture without Boundaries. Int. J. Comput. Vis. 67, 71–91. doi:10.1007/s11263-006-4068-8

Ma, Y., Kosecka, J., Soatto, S., Sastry, S., 2004. An Invitation to 3-D Vision: From Images to Geometric Models. Springer Science+Business Media, New York, USA. doi:10.1007/978-0-387-21779-6

Madjarova, V., Toyooka, S., Nagasawa, H., Kadono, H., 2003. Blooming processes in flowers studied by dynamic electronic speckle pattern interferometry (DESPI). Opt. Soc. Japan 10, 370–374.

Marinello, F., Pezzuolo, A., Gasparini, F., Sartori, L., 2013. Three-dimensional sensor for dynamic characterization of soil microrelief, in: Stafford, J. (Ed.), Precision Agriculture '13. Wageningen Academic Publishers, Wageningen, The Netherlands, pp. 71–78.

Martínez-Casasnovas, J.A., Ramos, M.C., Balasch, C., 2013. Precision analysis of the effect of ephemeral gully erosion on vine vigour using NDVI images. Precis. Agric. '13 777–783.

McCarthy, C.L., Hancock, N.H., Raine, S.R., 2010. Applied machine vision of plants: a review with implications for field deployment in automated farming operations. Intell. Serv. Robot. 3, 209–217. doi:10.1007/s11370-010-0075-2

Menesatti, P., Costa, C., Antonucci, F., Steri, R., Pallottino, F., Catillo, G., 2014. A low-cost stereovision system to estimate size and weight of live sheep. Comput. Electron. Agric. 103, 33–38. doi:10.1016/j.compag.2014.01.018

Möller, K., Jenz, M., Kroesen, M., Losert, D., Maurer, H.-P., Nieberg, D., Würschum, T., Ruckelshausen, A., 2016. Feldtaugliche Multisensorplattform für High-Throughput Getreidephänotypisierung - Aufbau und Datenhandling, 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. 137–140.

Moonrinta, J., Chaivivatrakul, S., Dailey, M.N., Ekpanyapong, M., 2010. Fruit detection, tracking, and 3D reconstruction for crop mapping and yield estimation. 2010 11th Int. Conf. Control Autom. Robot. Vis. 1181–1186. doi:10.1109/ICARCV.2010.5707436

Moreda, G.P., Ortiz-Cañavate, J., García-Ramos, F.J., Ruiz-Altisent, M., 2009. Non-destructive technologies for fruit and vegetable size determination – A review. J. Food Eng. 92, 119–136. doi:10.1016/j.jfoodeng.2008.11.004

Mousazadeh, H., 2013. A technical review on navigation systems of agricultural autonomous off-road vehicles. J. Terramechanics 50, 211–232. doi:10.1016/j.jterra.2013.03.004

Naio Technologies, 2014. Oz [WWW Document]. URL http://naio-technologies.com (accessed 2.10.16).

Nakarmi, A., Tang, L., 2012. Automatic inter-plant spacing sensing at early growth stages using a 3D vision sensor. Comput. Electron. Agric. 82, 23–31. doi:10.1016/j.compag.2011.12.011

Nayar, S.K., Nakagawa, Y., 1994. Shape from focus. IEEE Trans. Pattern Anal. Mach. Intell. 16, 824–831. doi:10.1109/34.308479

New Holland, 2010. IntelliFill system [WWW Document]. URL http://agriculture.newholland.com/ir/en/WNH/nhexcellence/Pages/IntelliFillSystem_detail.aspx (accessed 1.26.16).

Nissimov, S., Goldberger, J., Alchanatis, V., 2015. Obstacle detection in a greenhouse environment using the Kinect sensor. Comput. Electron. Agric. 113, 104–115. doi:10.1016/j.compag.2015.02.001

Noordam, J., Hemming, J., Van Heerde, C., Golbach, F., Van Soest, R., Wekking, E., 2005. Automated rose cutting in greenhouses with 3D vision and robotics : analysis of 3D vision techniques for stem detection, in: van Straten, G., Bot, G.P.., van Meurs, W.T.M., Marcelis, L.F.. (Eds.), Acta Horticulturae 691. ISHS, Leuven, Belgium, pp. 885–889.

odos imaging, 2014. real.iZ VS-1000 high-resolution time-of-flight [WWW Document]. URL http://www.odos-imaging.com/ (accessed 1.26.16).

Okutomi, M., Kanade, T., 1993. A multiple-baseline stereo. IEEE Trans. Pattern Anal. Mach. Intell. doi:10.1109/34.206955

Optimalog, 2014. Heliaphen [WWW Document]. URL http://www.optimalog.com/ (accessed 2.1.16).

Pallottino, F., Steri, R., Menesatti, P., Antonucci, F., Costa, C., Figorilli, S., Catillo, G., 2015. Comparison between manual and stereovision body traits measurements of Lipizzan horses. Comput. Electron. Agric. 118, 408–413. doi:10.1016/j.compag.2015.09.019

Phenospex, 2013. PlantEye [WWW Document]. URL https://phenospex.com/ (accessed

2.1.16).

Piron, A., van der Heijden, F., Destain, M.F., 2011. Weed detection in 3D images. Precis. Agric. 12, 607–622. doi:10.1007/s11119-010-9205-2

Piron, A., van der Heijden, F., Destain, M.F., 2010. Weed detection in 3D images. Precis. Agric. 12, 607–622. doi:10.1007/s11119-010-9205-2

Polder, G., Heijden, G.W.A.M. Van Der, Glasbey, C.A., Song, Y., Dieleman, J.A., 2009. Spy-See - Advanced vision system for phenotyping in greenhouses, in: MINET Conference: Measurement, Sensation and Cognition. London, UK, pp. 115–119.

Polder, G., Hofstee, J.W., 2014. Phenotyping large tomato plants in the greenhouse usig a 3D light-field camera. ASABE CSBE/SCGAB Annu. Int. Meet. 1, 153–159.

Polder, G., Lensink, D., Veldhuisen, B., 2013. PhenoBot - a robot system for phenotyping large tomato plants in the greenhouse using a 3D light field camera, in: Phenodays. Wageningen UR, Vaals, Holand.

Reina, G., Milella, A., 2012. Towards autonomous agriculture: automatic ground detection using trinocular stereovision. Sensors (Basel). 12, 12405–23. doi:10.3390/s120912405

Reina, G., Milella, A., Nielsen, M., Worst, R., Blas, M.R., 2015. Ambient awareness for agricultural robotic vehicles. Biosyst. Eng. 1–19. doi:10.1016/j.biosystemseng.2015.12.010

Rosell, J.R., Sanz, R., 2012. A review of methods and applications of the geometric characterization of tree crops in agricultural activities. Comput. Electron. Agric. 81, 124–141. doi:10.1016/j.compag.2011.09.007

Rovira-Más, F., Han, S., Wei, J., Reid, J.F., 2007. Autonomous guidance of a corn harvester using stereo vision. Agric. Eng. Int. CIGR Ejournal IX, 1–13.

Rovira-Más, F., Zhang, Q., Reid, J.F., 2005. Creation of three-dimensional crop maps based on aerial stereoimages. Biosyst. Eng. 90, 251–259. doi:10.1016/j.biosystemseng.2004.11.013

Saeys, W., Lenaerts, B., Craessaerts, G., De Baerdemaeker, J., 2009. Estimation of the crop density of small grains using LiDAR sensors. Biosyst. Eng. 102, 22–30. doi:10.1016/j.biosystemseng.2008.10.003

Salvi, J., Fernandez, S., Pribanic, T., Llado, X., 2010. A state of the art in structured light patterns for surface profilometry. Pattern Recognit. 43, 2666–2680. doi:10.1016/j.patcog.2010.03.004

Santos, T.T., Oliveira, A.A. De, 2012. Image-based 3D digitizing for plant architecture analysis and phenotyping, in: XXV Conference on Graphics, Patterns and Images. Ouro Preto, Brazil, pp. 21–28.

Savarese, S., 2005. Shape Reconstruction from Shadows and Reflections. California Institute

of Technology.

Scharstein, D., Szeliski, R., 2002. A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms. Int. J. Comput. Vis. 47, 7–42.

Schwarte, R., Heinol, H., Buxbaum, B., Ringbeck, T., 1999. Principles of Three-Dimensional Imaging Techniques, in: Jähne, B., Haußecker, H., Greißler, P. (Eds.), Handbook of Computer Vision and Applications. Academic Press, Heidelberg, Germany, pp. 464–482.

Šeatović, D., Kuttere, H., Anken, T., Holpp, M., 2009. Automatic weed detection in grassland, in: 67th International Conference on Agricultural Engineering. Hanover, Germany, pp. 187–192.

Seitz, S., Curless, B., Diebel, J., Scharstein, D., Szeliski, R., 2006. A Comparison and Evaluation of Multi-View Stereo Reconstruction Algorithms, in: IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE, New York, USA, pp. 519–528. doi:10.1109/CVPR.2006.19

Shalal, N., Low, T., Mccarthy, C., Hancock, N., 2013. A review of autonomous navigation systems in agricultural environments, in: Innovative Agricultural Technologies for a Sustainable Future. Society for Engineering in Agriculture (SEAg), Barton, Australia, pp. 1–16.

Strothmann, W., Ruckelshausen, A., Hertzberg, J., 2014. Multiwavelenght laser line profile sensing for agricultural crop, in: Berghmans, F., Mignani, A.G., De Moor, P. (Eds.), Proceedings SPIE 9141 Optical Sensing and Detection III. SPIE, Brussels, Belgium. doi:10.1117/12.2072122

Sun, C., 2002. Fast optical flow using 3D shortest path techniques. Image Vis. Comput. 20, 981–991.

Tabb, A., 2013. Shape from Silhouette probability maps: reconstruction of thin objects in the presence of silhouette extraction and calibration error. 2013 IEEE Conf. Comput. Vis. Pattern Recognit. 161–168. doi:10.1109/CVPR.2013.28

Tanaka, T., Kataoka, T., Ogura, H., Shibata, Y., 2014. Evaluation of rotary tillage performance using resin-made blade by 3D-printer, in: 18th International Conference of the ISTVS. Seul, Korea.

The University of Sidney, 2015. Ladybird [WWW Document]. URL http://www.acfr.usyd.edu.au/ (accessed 2.1.16).

Thilakarathne, B.L.S., Rajagopalan, U.M., Kadono, H., Yonekura, T., 2014. An optical interferometric technique for assessing ozone induced damage and recovery under cumulative exposures for a Japanese rice cultivar. Springerplus 3, 89. doi:10.1186/2193-1801-3-89

Tiziani, H.J., 1989. Optische Methoden der 3-D-Messtechnik und Bildverarbeitung, in: Ahlers, Rolf-Jürgen (Hrsg.): Bildverarbeitung: Forschen, Entwickeln, Anwenden. Ostfildern, Germany, pp. 1–26.

UMR Itap, 2015. Becam [WWW Document]. URL http://itap.irstea.fr/ (accessed 2.1.16).

Underwood, J., Calleija, M., Nieto, J., Sukkarieh, S., 2013. A robot amongst the herd: remote detection and tracking of cows, in: Ingram, L., Cronin, G., Sutton, L. (Eds.), Proceedings of the 4th Australian and New Zealand Spatially Enabled Livestock Management Symposium. Faculty of Agriculture and Environment, and the Faculty of Veterinary Science, The University of Sydney, Camden, Australia, p. 52.

UNEP, 2010. Assessing the Environmental Impacts of Consumption and Production: Priority Products and Materials, A Report of the Working Group on the Environmental Impacts of Products and Materials to the International Panel for Sustainable Resource Management.

Van der Stuyft, E., Schofield, C.P., Randall, J.M., Wambacq, P., Goedseels, V., 1991. Development and application of computer vision systems for use in livestock production. Comput. Electron. Agric. 6, 243–265. doi:10.1016/0168-1699(91)90006-U

Vázquez-Arellano, M., Reiser, D., Garrido, M., Griepentrog, H.W., 2016. 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.

Velodyne, 2015. Puck VLP-16 [WWW Document]. URL http://velodynelidar.com/ (accessed 12.3.15).

Viazzi, S., Bahr, C., Van Hertem, T., Schlageter-Tello, A., Romanini, C.E.B., Halachmi, I., Lokhorst, C., Berckmans, D., 2014. Comparison of a three-dimensional and two-dimensional camera system for automated measurement of back posture in dairy cows. Comput. Electron. Agric. 100, 139–147. doi:10.1016/j.compag.2013.11.005

Vision Robotics, 2014. Proto prune [WWW Document]. URL http://www.visionrobotics.com/ (accessed 1.26.16).

Vos, J., Marcelis, L., De Visser, P., Struik, P., Evers, J., 2007. Functional–structural plant modelling in crop production, Journal of experimental botany. Wageningen, The Netherlands. doi:10.1093/jxb/erp345

Wang, Q., Nuske, S., Bergerman, M., Singh, S., 2013. Experimental Robotics: The 13th International Symposium on Experimental Robotics, in: Desai, P.J., Dudek, G., Khatib, O., Kumar, V. (Eds.), . Springer International Publishing, Heidelberg, pp. 745–758. doi:10.1007/978-3-319-00065-7_50

Wang, Q., Nuske, S., Bergerman, M., Singh, S., 2012. Automated crop yield estimation for apple orchards, in: International Symposium on Experimental Robotics. Quebec City, Canada.

Wang, Q., Zhang, Q., Rovira-Más, F., Tian, L., 2011. Stereovision-based lateral offset measurement for vehicle navigation in cultivated stubble fields. Biosyst. Eng. 109, 258–265. doi:10.1016/j.biosystemseng.2011.04.006

Wei, J., Reid, J.F., Han, S., 2005. Obstacle Detection Using Stereo Vision To Enhance Safety of Autonomous Machines. Trans. Am. Soc. Agric. Eng. 48, 2389–2397.

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

Weiss, U., Biber, P., Laible, S., Bohlmann, K., Zell, A., Gmbh, R.B., 2010. Plant Species Classification using a 3D LIDAR Sensor and Machine Learning, in: Machine Learning and Applications. IEEE, Washington, D.C., USA, pp. 339–345.

Westoby, M.J., Brasington, J., Glasser, N.F., Hambrey, M.J., Reynolds, J.M., 2012. "Structure-from-Motion" photogrammetry: A low-cost, effective tool for geoscience applications. Geomorphology 179, 300–314. doi:10.1016/j.geomorph.2012.08.021

Wolff, A., 2012. Phänotypisierung in Feldbeständen mittels 3D-Lichtschnitt-Technik. Strube Research GmbH, Söllingen, Germany.

Woodham, R., 1980. Photometric method for determining surface orientation from multiple images. Opt. Eng. 19, 139–144.

Wu, J., Tillett, R., McFarlane, N., Ju, X., Siebert, J.P., Schofield, P., 2004. Extracting the three-dimensional shape of live pigs using stereo photogrammetry. Comput. Electron. Agric. 44, 203–222. doi:10.1016/j.compag.2004.05.003

Wulder, M. a., White, J.C., Nelson, R.F., Næsset, E., Ørka, H.O., Coops, N.C., Hilker, T., Bater, C.W., Gobakken, T., 2012. Lidar sampling for large-area forest characterization: A review. Remote Sens. Environ. 121, 196–209. doi:10.1016/j.rse.2012.02.001

Yang, L., Noguchi, N., 2012. Human detection for a robot tractor using omni-directional stereo vision. Comput. Electron. Agric. 89, 116–125. doi:10.1016/j.compag.2012.08.011

Yin, X., Noguchi, N., Choi, J., 2013. Development of a target recognition and following system for a field robot. Comput. Electron. Agric. 98, 17–24. doi:10.1016/j.compag.2013.07.005

Zarco-Tejada, P.J., Diaz-Varela, R., Angileri, V., Loudjani, P., 2014. Tree height quantification using very high resolution imagery acquired from an unmanned aerial vehicle (UAV) and automatic 3D photo-reconstruction methods. Eur. J. Agron. 55, 89–99. doi:10.1016/j.eja.2014.01.004

Zhao, Y., Sun, Y., Cai, X., Liu, H., Lammers, P.S., 2012. Identify Plant Drought Stress by 3D-

Based Image. J. Integr. Agric. 11, 1207–1211. doi:10.1016/S2095-3119(12)60116-6

Zion, B., 2012. The use of computer vision technologies in aquaculture – A review. Comput. Electron. Agric. 88, 125–132. doi:10.1016/j.compag.2012.07.010

3 Part II: Reconstruction of geo-referenced maize plants using a consumer time-of-flight camera in different agricultural environments

Abstract

Crop phenotyping is a prerequisite to enable robots doing agricultural tasks, evaluating crop status for farm management, and relating genotypes to phenotypes for crop breeding among others. Optical three dimensional (3-D) sensors have been preferred since they provide more information about the complex plant architecture. The improvement of time-of-flight (TOF) cameras together with their reduced economical costs have provided an appropriate tool for tasks that require detailed information of the agricultural environment. In this paper, 3-D reconstruction of maize is performed in different environments, from controlled greenhouse to the open field, to evaluate the capabilities of a consumer camera.

Keywords: 3-D sensors, time-of-flight, agricultural automation, plant phenotyping

3.1 Introduction

Crop phenotyping is a prerequisite to enable robots doing agricultural tasks, evaluating crop status for farm management (Griepentrog et al., 2010), and relating genotypes to phenotypes for crop breeding among others; yet it remains a bottleneck (Furbank and Tester, 2011) due to the time-consuming measuring methodologies and systems. Moreover, it is also important that the acquired sensor data is accurately and precisely geo-referenced to know the position in space of every plant and plant element. Advances in off-the-shelf 3-D vision sensors are opening new possibilities since they provide more information compared with two dimensional (2-D) sensors in a cost-effective manner; however, it comes at the cost of more computer power and data handling.

The Kinect v2 is an example of a CTC that has appealing characteristics like: high depth image pixel resolution, NIR stream for night vision, and a relative robustness against sunlight. For geo-referencing optical information, real time kinematic-global navigation satellite system (RTK-GNSS) is limited to outdoor conditions and its accuracy (centimetre-level) is not better

than the (sub-centimetre-level) robotic total station; therefore, the latter was used for this research.

The aim of this research is to present a methodology for reconstructing maize plants using a CTC mounted on a field robot. This robot navigates in different agricultural scenarios using a robotic total station for geo-referencing the position of the CTC, and thus, the generated point clouds.

3.2 Materials and Methods

A robotic platform, depicted in Figure 3-1, developed at the University of Hohenheim, was used for data acquisition. The dimensions of the robotic platform are: length 600 mm, width 500 m, height 1100 mm. The vehicle carries a CTC for data acquisition mounted on an extruded aluminium frame. The robotic platform software was developed using the Robot Operating System (ROS Indigo), an open source middleware running on Linux (Ubuntu 14.04), and programmed in a combination of C++ and Python programming languages. For fast calibration, point measurement, and importing data from the total station into ROS; the Trimble SCS900 Site Controller (Software Version 3.4.0) was used. The prism position data was time stamped and helped to refer the transforms to the global frame (Reiser et al., 2016a).
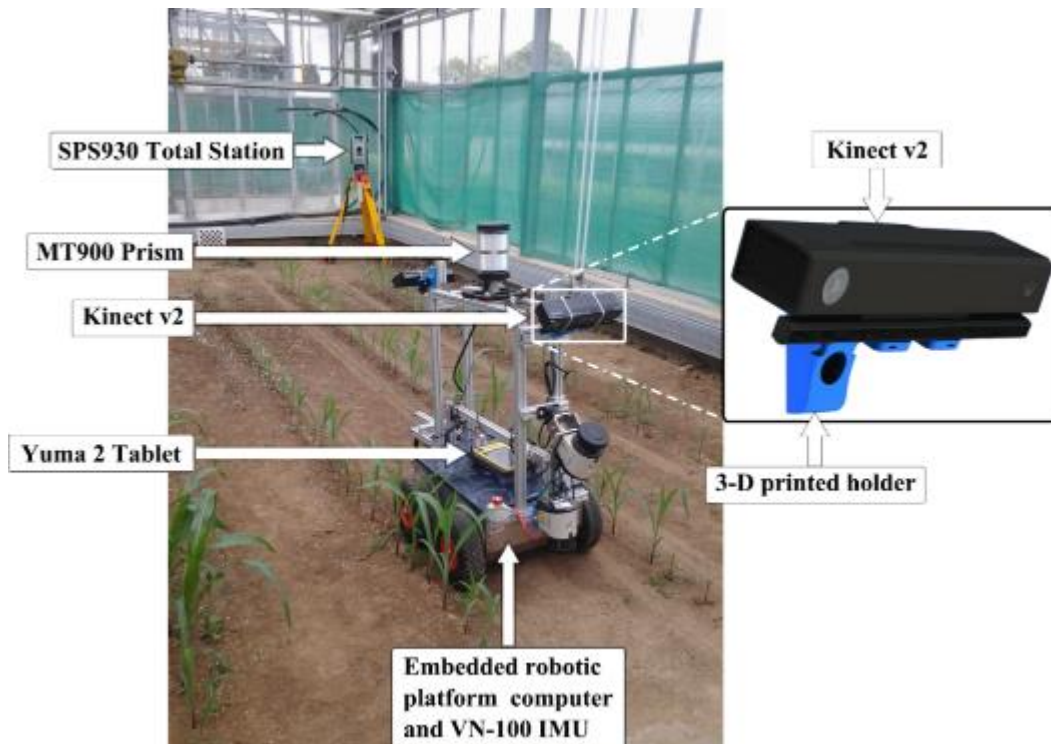
Figure 3-1: Robotic platform for 3-D data acquisition in a greenhouse and the utilized total station.

The SPS930 robotic total station (Trimble Navigation Limited, Sunnyvale, USA) was used to track the precise 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 1.07m. The total station data was sent to a Yuma 2 Tablet Computer (Trimble, Sunnyvale, USA). The CTC has 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 30°. The CTC outputs three image streams: a depth image stream of 512 x 424 pixels, a NIR stream of 512x424 pixels, and a colour stream of 1920x1080 pixels. Depth, infrared, and Red-Green-Blue (RGB) images were acquired from 23.04.2015 to 01.07.2015, and a total of 9 tests were done driving the robot through the tracks using a remote joystick at a constant speed of *circa* 0.05 m s^{-1}. Every track was passed two times starting from each side.

3.3   Results and Discussion

A first trail was done using two CTCs mounted on the robot, one pointing forwards and the other backwards, in order to have two different perspectives and to avoid passing two times through the same track. However, the high amount of data simultaneously coming from the two CTCs rapidly overloaded the computer acquisition system. Therefore, it was decided to use only one CTC and drive two times through the same track from opposite sides. Due to the high spectral reflectance of plants in the *NIR plateau* (737-1000 nm), it was possible to obtain depth images of maize in different lighting conditions as shown in Figure 3-2.
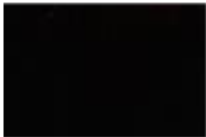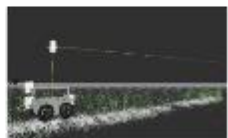
Figure 3-2:   CTC output in different environments, lighting conditions, and maize heights (*mean height$_{greenhouse\ sun-shadow}$*=98 mm, *mean height$_{greenhouse\ night}$*=50 mm, *mean height$_{open\ field\ sunny}$*=500 mm).

Preliminary results (Figure 3-3) show that it is possible to reconstruct maize assembling point clouds in different agricultural environments and light conditions. At night, the maize 3-D reconstruction was possible with the least amount of noise; inside the greenhouse, the light variability was a source of noise; and in the open field, most of the depth image (excluding maize plants) was saturated by noise. The wind conditions were favourable during the data acquisition in the open field, and if they were not, a wind protection could have been used- as most of the robotic phenotyping platforms do.
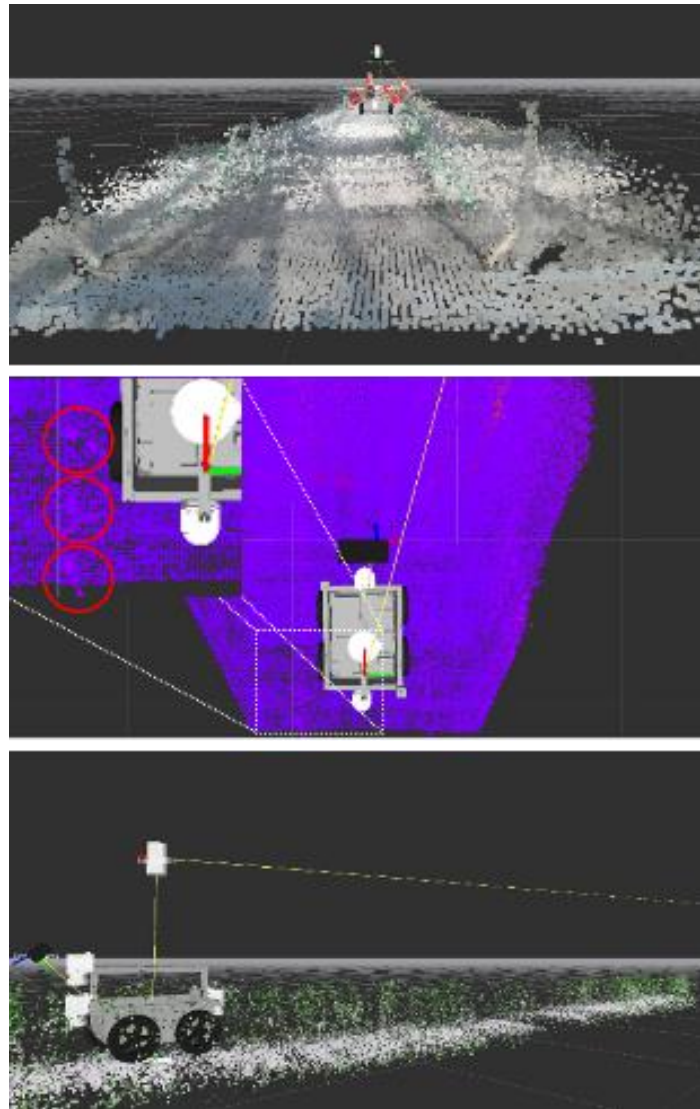
Figure 3-3:   Point cloud reconstruction of maize in greenhouse in sun-shadow (top), greenhouse at night (middle) with the maize inside the circles, and open field in a cloudless sunny day (down).

3.4 Conclusions

The CTC used in this research has a lot of potential in agricultural applications mainly due to the capability of providing depth information under different lighting conditions. Although it was designed for other purposes, it has shown that it can stream depth information under different environments, and even though it does not perform well under sunlight, it is still possible to obtain depth data of maize. Better results would be expected if a shadowing device is used. Surface reconstruction algorithms like Kintinuous could be also applied to the point clouds for a better representation of the leaf surface; however, it was out of the objectives of this research.

References

Furbank, R.T., Tester, M., 2011. Phenomics--technologies to relieve the phenotyping bottleneck. Trends Plant Sci. 16, 635–44. doi:10.1016/j.tplants.2011.09.005

Griepentrog, H.W., Ruckelshausen, A., Joergensen, R.N., Lund, I., 2010. Autonomous systems for plant protection, in: Oerke, E.C., Gerhards, R., Menz, G., Sikora, R.A. (Eds.), Precision Crop Protection - the Challenge and Use of Heterogeneity. Springer, pp. 323–334.

Reiser, D., Miguel, G., Arellano, M.V., Griepentrog, H.W., Paraforos, D.S., 2016. Crop row detection in maize for developing navigation algorithms under changing plant growth stages. Adv. Intell. Syst. Comput. 417, 371–382. doi:10.1007/978-3-319-27146-0_29

4   Part III: 3-D reconstruction of maize plants using a time-of-flight camera

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 algorithm. The incoming point cloud was previously filtered using the Random Sample Consensus 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 camera, compared with the ones from two light detection and ranging 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.

Keywords: 3-D sensors; Kinect v2; time-of-flight; agricultural robotics; precision farming

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^{robot}_{tof} Transformation matrix from the TOF sensor to the robotic platform coordinate system.

T^{ts}_{robot} 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.

4.1 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 commercial TOF cameras and their suitability for agricultural applications (Kazmi et al., 2014; Klose et al., 2009). 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 et al., 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) 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 x 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.

4.2 Materials and Methods

4.2.1 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 (Figure 3-1). 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 4-1 shows a more detailed explanation.

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 4-1: General specifications of the computer hardware, their operating systems and installed software.

The basic principle for optical depth measurement behind the Kinect v2 sensor is time-of-flight, specifically continuous wave modulation TOF (Elise Lachat et al., 2015). 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 short-ened 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 4-2. 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 Figure 4-1).
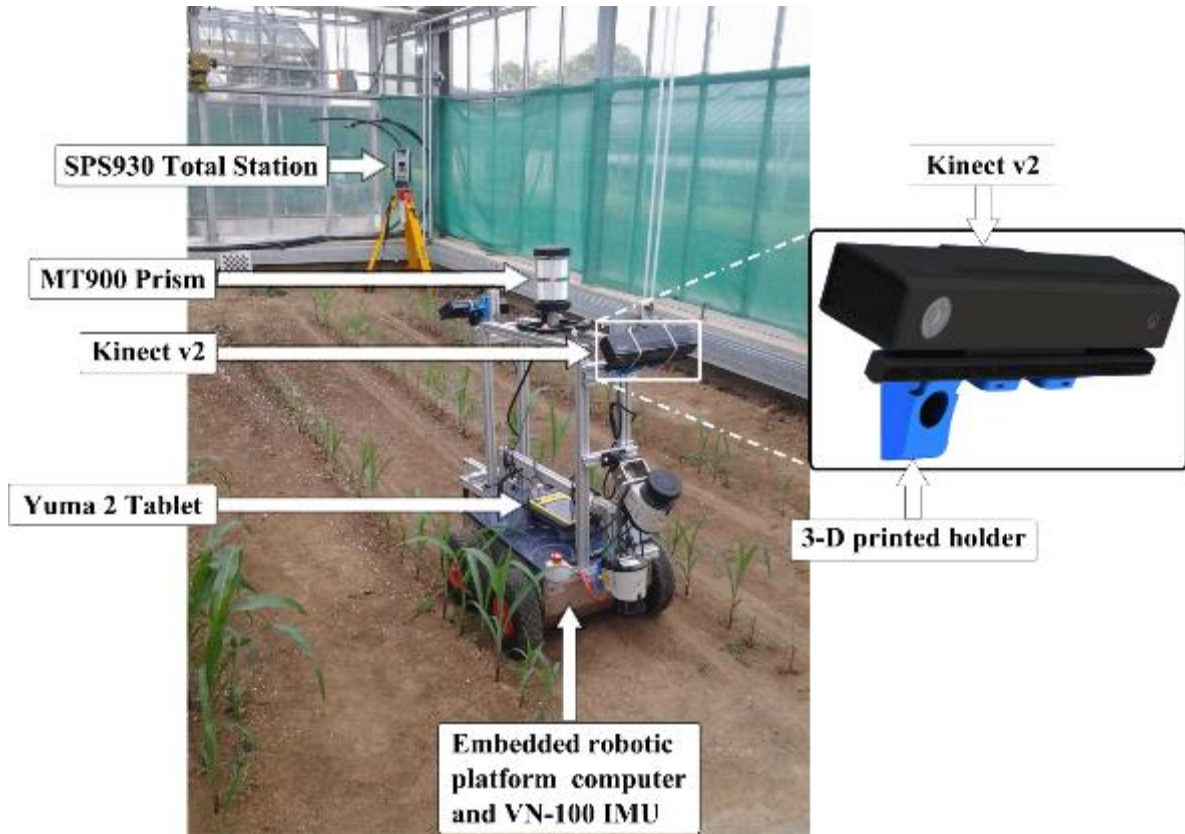
Figure 4-1:   Acquisition system for the experiment depicting the TALOS robotic platform carrying multiple sensors. For this research paper, the data from the Kinect v2, VN-100 IMU and the geo-referenciation system were used. The geo-referenciation system consists of the SPS930 total station, MT900 target prism and Yuma 2 Tablet.

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 ToolboxTM of MATLAB R2016b was used as an interface for exporting the stitched point cloud from ROS to MATLAB for further segmentation and visual representation.

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 to 4.5m
	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 to 850 nm
Data transmission	Interface standard	USB 3.0

Table 4-2: General specifications of the TOF camera (Microsoft, 2017).

4.2.2   Experimental setup

The experiment was conducted in a greenhouse of 21 m2 (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 hour 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 minutes, when the measurements vary ±1 mm (E. Lachat et al., 2015). 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., 2016b). The data was then acquired by driving the robot through the 4 paths (5 crop rows) using a remote joystick at a constant speed of circa 0.05 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 to 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 Figure 4-2); 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.
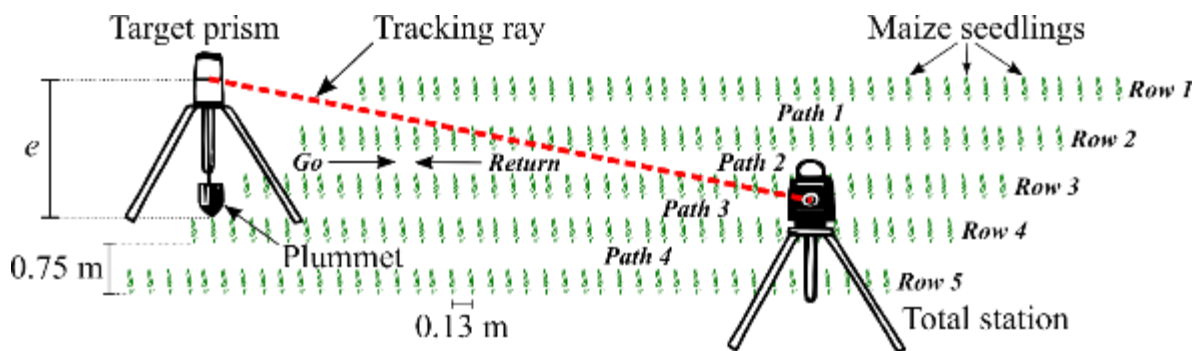
Figure 4-2:   Arrangement of seedling position measurement. The distance e is a constant value between the target prism and the tip of the plummet. The path and crop row numbering as well as the driving direction "go" and "return", and the seedling spacing are also depicted.

4.2.3 Data processing

4.2.3.1 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.

4.2.3.2 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 Figure 4-3.
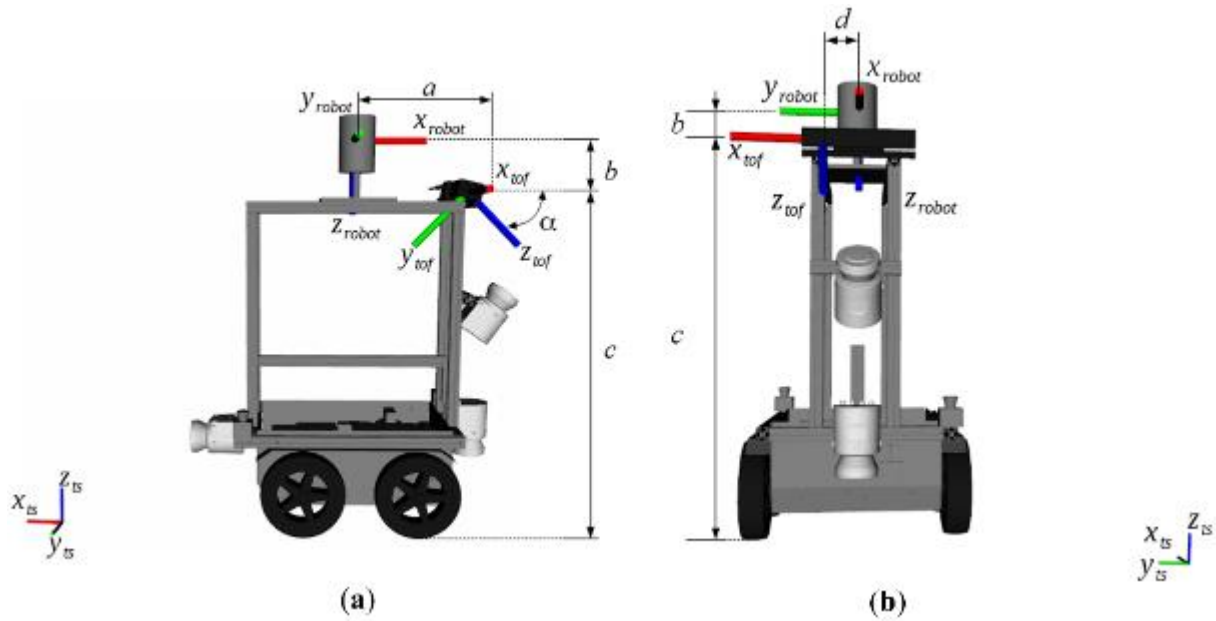
Figure 4-3:   Schematic Representation of the used robotic platform showing the coordinate frames of the TOF camera $(x_{tof}\ y_{tof}\ z_{tof})$, robotic platform $(x_{robot}\ y_{robot}\ z_{robot})$ and total station $(x_{ts}\ y_{ts}\ z_{ts})$ and their locations relative to each other. (a) Side view; (b) front view.

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 Equation 4-1.

$$[x_{robot}\ y_{robot}\ z_{robot}\ 1]^T = [T_{tof}^{robot}] \times [x_{tof}\ y_{tof}\ z_{tof}\ 1]^T \qquad \text{(Equation 4-1)}$$

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 $Rot\left('yaw',-\frac{\pi}{2}\right)$, a rotation $Rot\left('roll',\frac{\pi}{2}-\alpha\right)$ and three translations $(Trans(x,+a),\ Trans(z,+b)$ and $Trans(y,-d))$. The complete transformation is represented in Equation 4-2.

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

(Equation 4-2)

The general matrix form of all rotations and translations are shown in equation 4-3, and by substituting them in Equation 4-2 and 4-5, a matrix form representation of the equations can be generated.

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

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

(Equation 4-3)

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

Then, the robot frame needs to be transformed to the total station frame as depicted in Equation 4-4

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

(Equation 4-4)

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 Equation 3-5.

$$T_{robot}^{ts} = \underbrace{Rot('pitch', -\pi)}_{coordinate\ system\ matching}$$
$$\cdot Rot('pitch', -\theta) \cdot Rot('roll', -\varphi) \cdot Rot('yaw', -\psi)$$
$$\cdot Trans(x, \pm x_{ts}) \cdot Trans(y, -y_{ts}) \cdot Trans(z, -z_{ts})$$

(Equation 4-5)

4.2.3.3 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 ToolboxTM 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 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 Figure 4-4. 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.
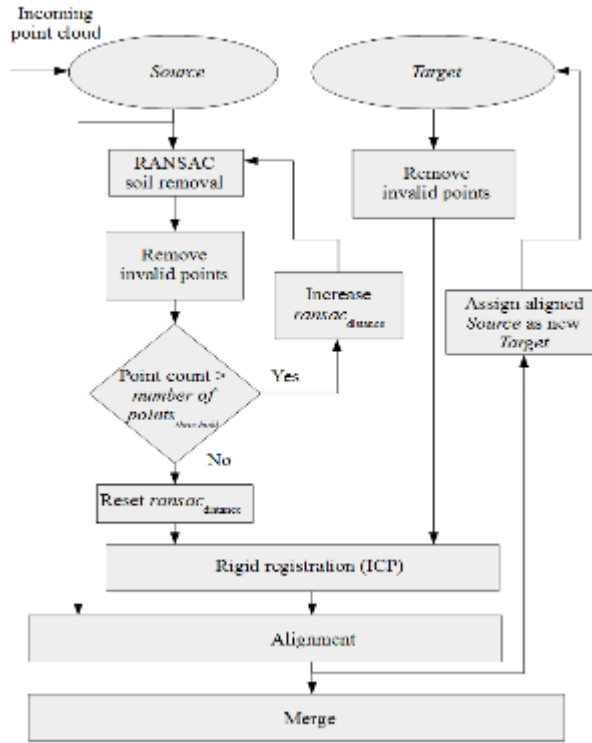
Figure 4-4:   Complete workflow of the adaptive point cloud rigid registration and stitching process.

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 Equation 4-6:

$$number\ of\ points_{threshold} = \frac{number\ of\ points_{source}}{2 \times plant\ points\%} \qquad \text{(Equation 4-6)}$$

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 equation 6, the resulting number of points of the threshold point cloud was equal to ~80,000 points. Once the RANSAC 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.

4.2.3.4 Point cloud segmentation methodology

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

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.

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.

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.

4.2.3.5 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:

− Project a parallel plane with a normal distance of 50 mm above the RANSAC plane model.

− 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.

− 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.

− 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.

4.3   Results and Discussion

After the point cloud paths were stitched in ROS, they were saved in ".bag" files. Then, the Robotics System ToolboxTM 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.

4.3.1   Point cloud segmentation

The result of the registration and stitching of all point clouds acquired along path 2 is shown in Figure 4-5a while going and Figure 4-5b while returning. It can be seen that due to the complex

plant architecture, the 3-D shape reconstruction is slightly different when driving in the opposite direction; but still the maize leaves are clearly defined.
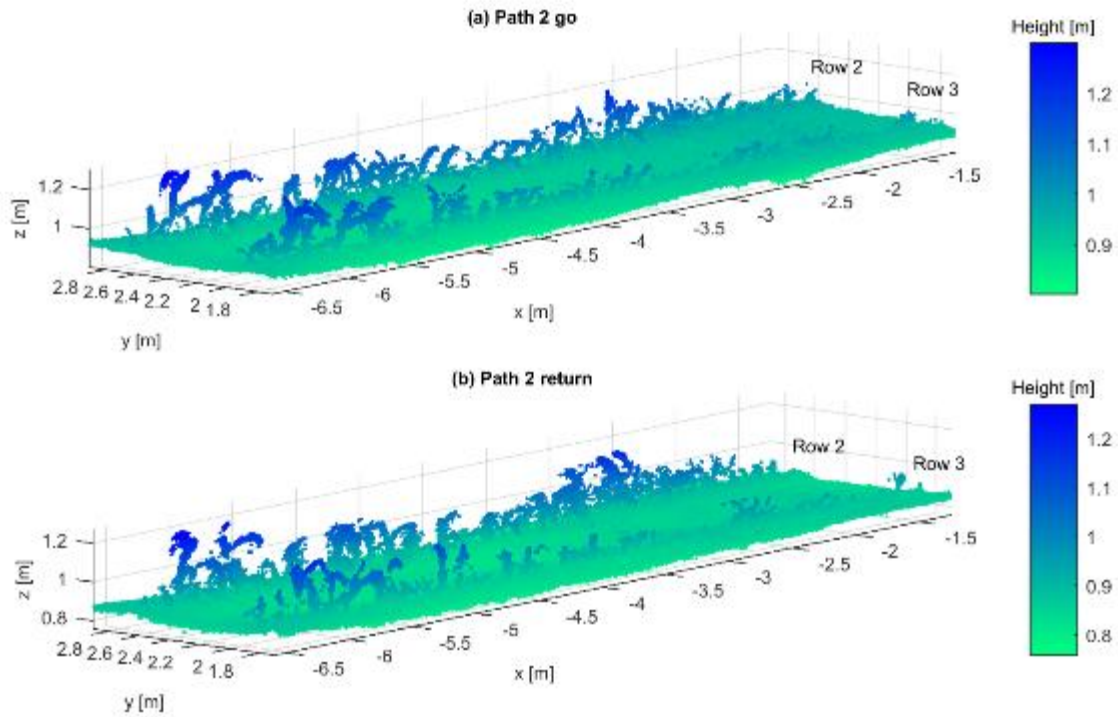
Figure 4-5:   Complete path 2 registration and stitching where the crop rows 2 and 3 are clearly recognisable (a) go; and (b) return.

After the registration and stitching process, the resulting 3-D point cloud of every path is presented in Table 4-3 for both directions (go and return). The raw data column represents the reconstructed path containing two crop rows as seen in Figure 4-3. 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 3-3, the average point cloud count of a scene, such as the ones depicted in Figure 4-5, is roughly 10 million points where just 76,064 belonged to plants.

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 4-3: Point cloud reduction from path point clouds for plant point extraction.

By dividing the path into two halves, in order to obtain the individual crop rows, the results presented in Table 4-4 are comparable with the ones of Table 4-3. 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 Figure 4-1 and Figure 4-2. 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.

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%)

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

In the paper by Garrido et al. (2015), the vertical LIDAR was mounted in an advantageous position (see again Figure 4-1 and Figure 4-2), 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 3-4), 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. Figure 4-6 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 (Figure 4-6c, and Figure 4-6e). 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 Figure 4-6d (e.g. plant 11) and Figure 4-7c (e.g. plant 3) seem more defined compared with the ones in Figure 4-6f and Figure 4-6c, 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 recognize the plants.
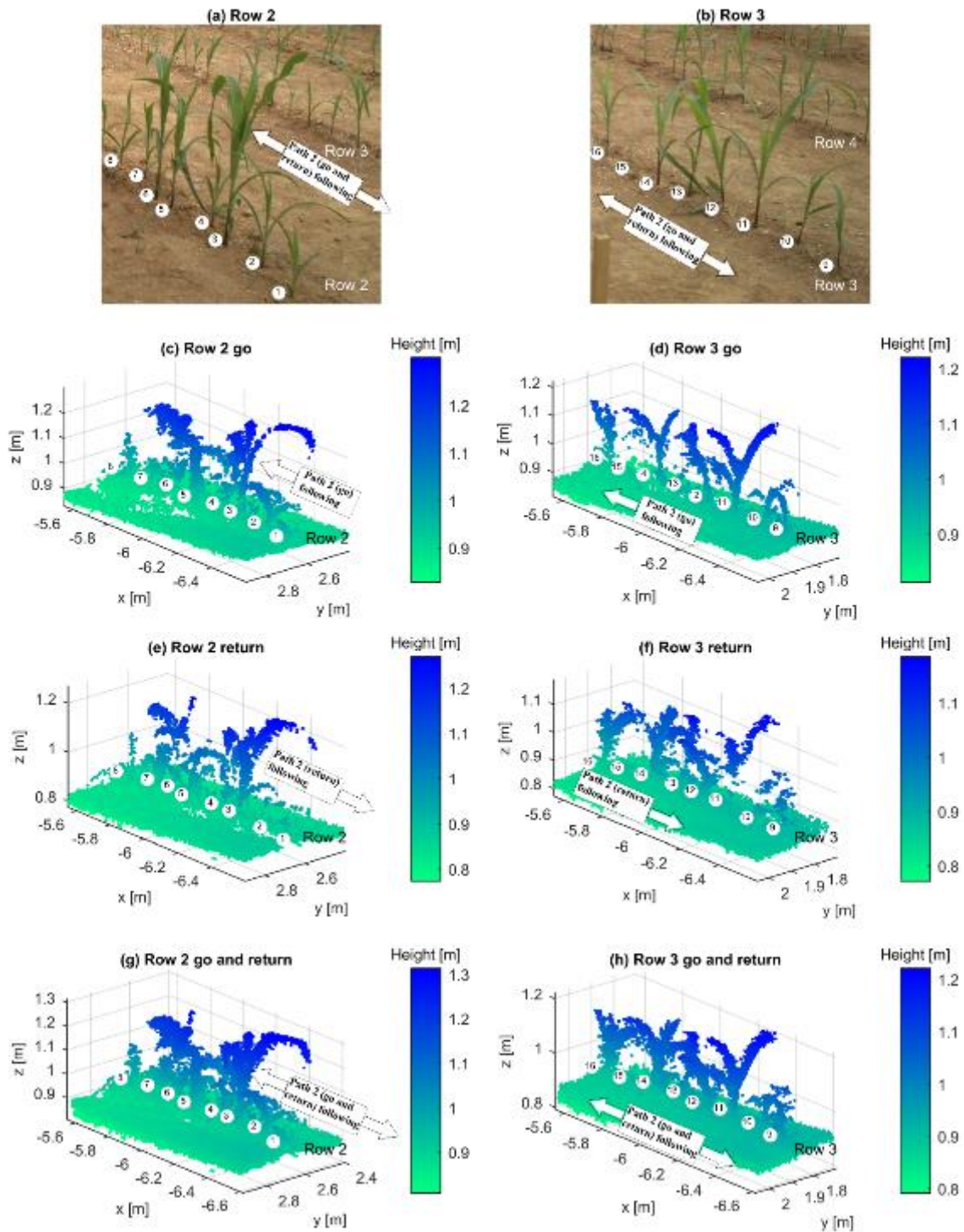
Figure 4-6:   Path 2 section. (a) row 2 RGB image; (b) row 3 RGB image; (c) row 2 go; (d) row 3 go; (e) row 3 return; (f) row 3 return; (g) row 2 fusion; (h) row 3 fusion.

It is also interesting to compare the point clouds while going, returning and the fusion of both of them. In Figure 4-6e and Figure 4-6f, 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 Figure 4-6c and Figure 4-6d, 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.
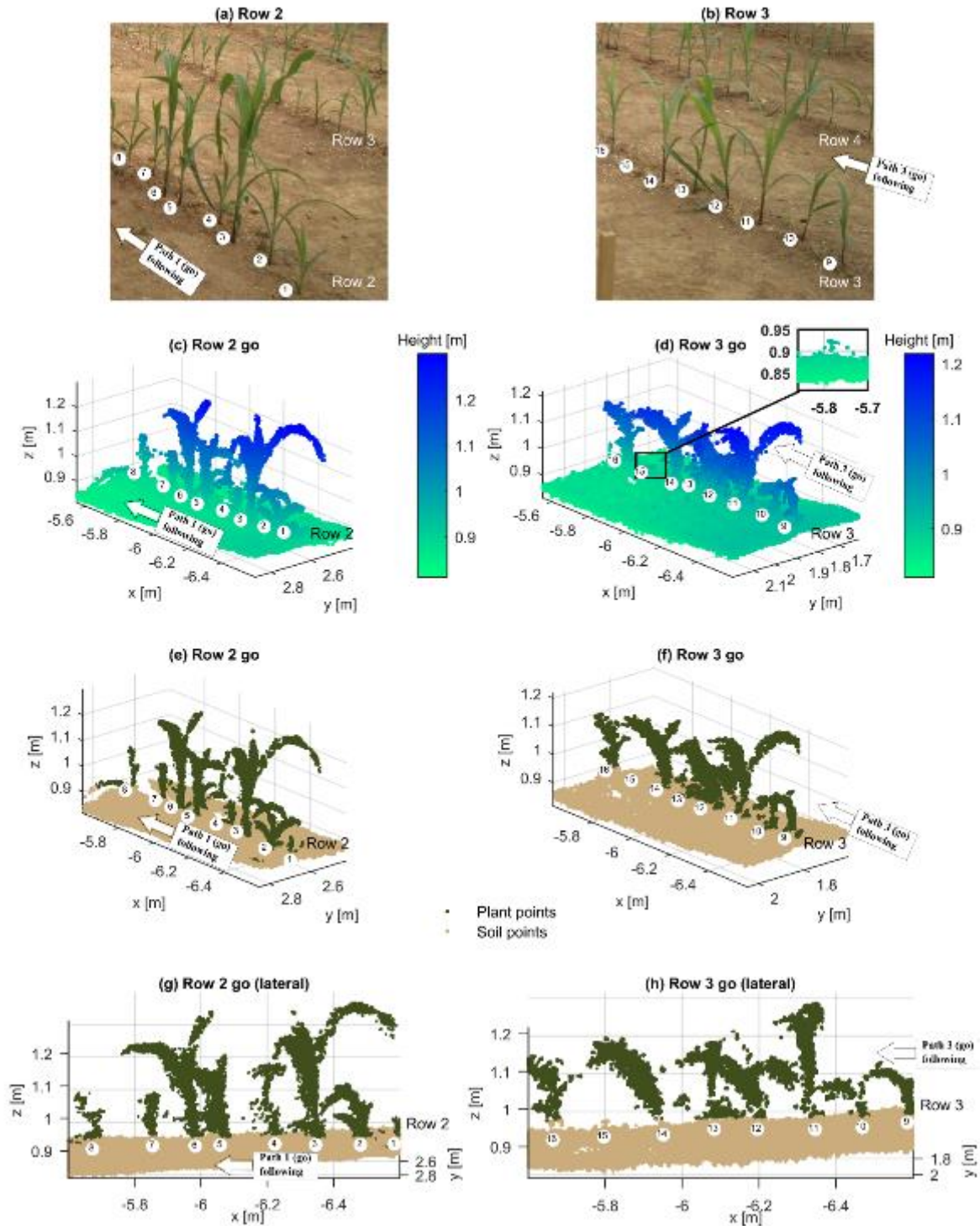
Figure 4-7:   Path 1 section. (a) row 2 RGB image; (b) row 3 RGB image; (c) row 2 go; (d) row 3 go; (e) row 3 return; (f) row 3 return (g) row 2 return lateral (h) row 3 return lateral.

The advantage of the TOF light volume technique is that some information can be obtained from small and thin objects. In Figure 4-7b, 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 Figure 4-7f and Figure 4-7h, 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 4-5 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. Figure 4-8b depicts the reconstruction of the whole field (Figure 4-8a) going and returning with 246,182 plant and 23,832,277 soil points.

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%)	

Table 4-5: Plant point cloud fusion and down sample.
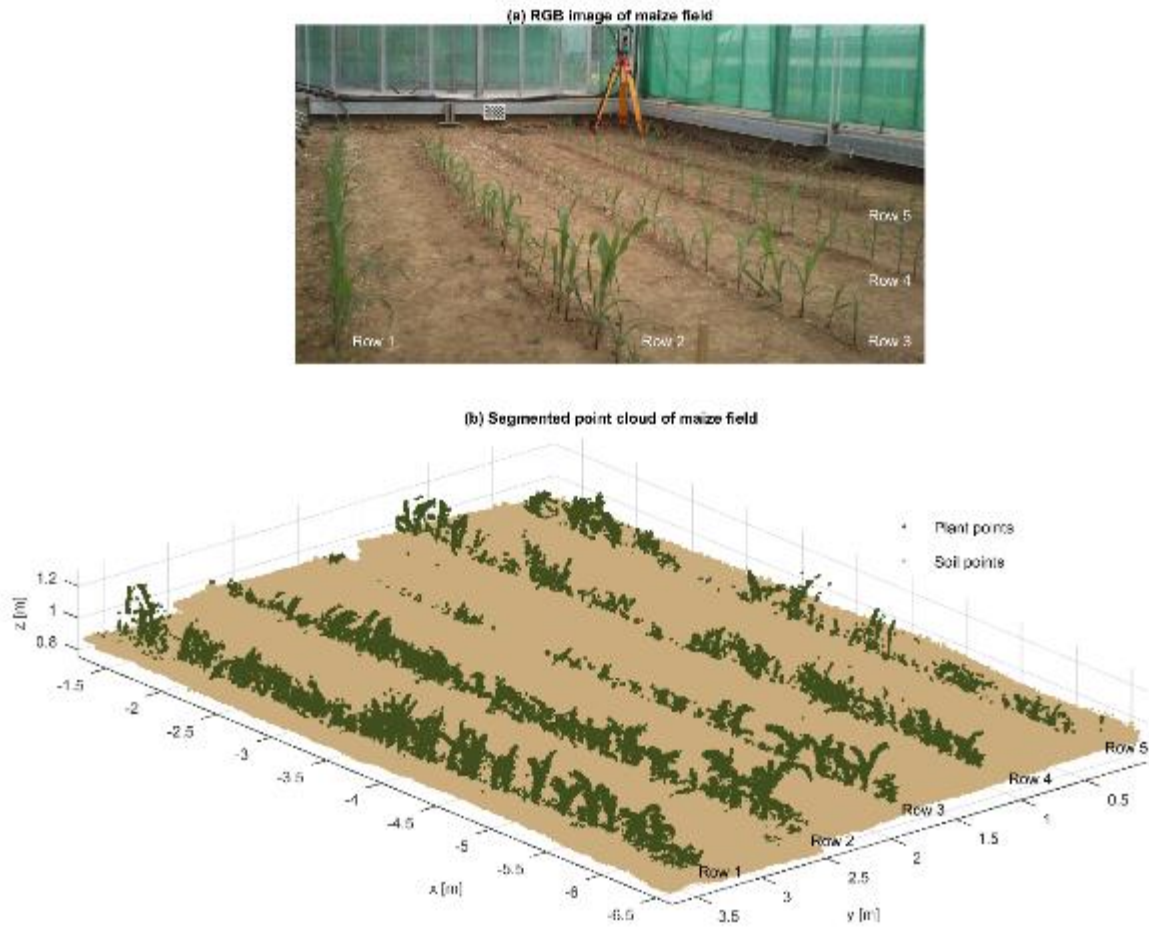
Figure 4-8:   Maize field depicted as: (a) RGB image and (b) down sampled and segmented
                3-D reconstruction going and returning.

4.3.2   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 ham-
mered 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 reconstruction,
but also to recalibrate the registration and stitching algorithm in the case that it starts accumu-
lating 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 Figure 4-9a) 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 Figure 4-9b. Then, a 2-D k-means clustering was applied to those points. Since some plants were not detected because of reasons such 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 (Figure 4-7g and Figure 4-7h). 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.
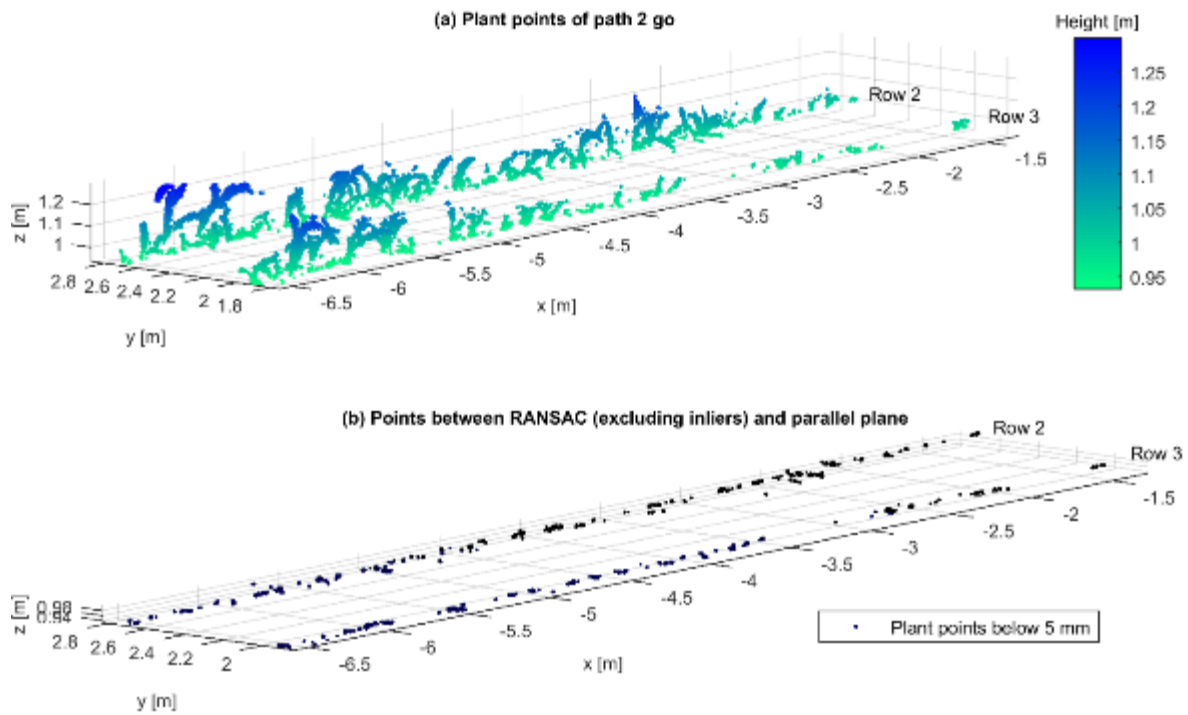
Figure 4-9:   Point cloud segmentation; (a) showing the plants after applying the RANSAC soil removal; (b) plant point cloud between RANSAC and parallel plane.

The alignment in the y axis is evident in Figure 4-10, 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.
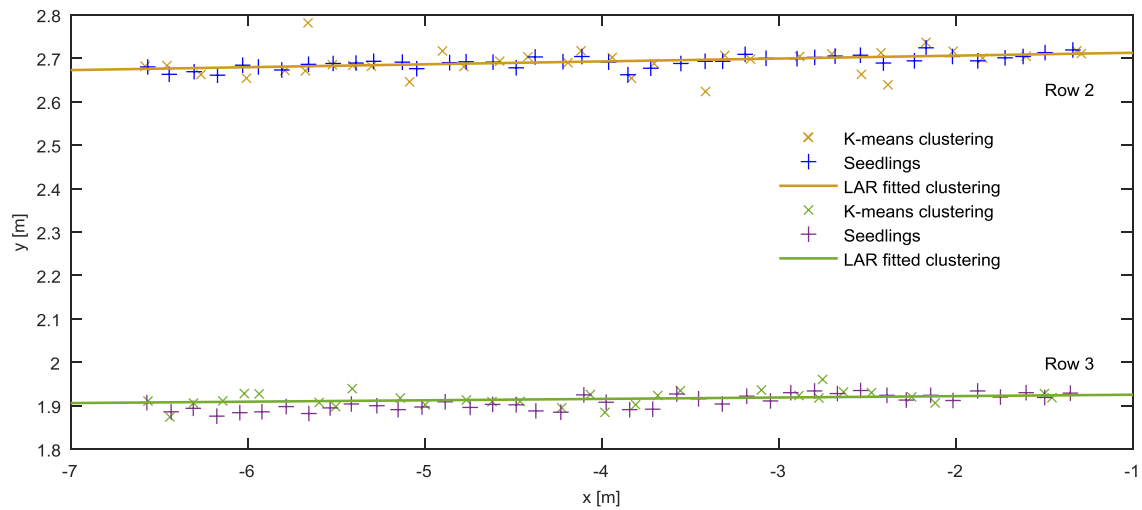
Figure 4-10:   Validation methodology showing the k-means clustering result (x) with a LAR fitting curve of the clustering and the seedling locations (+).

In order to evaluate the accuracy of the clustering in Figure 4-10, the position in the x and y coordinates was compared with the ground truth. The mean and the standard deviation are shown in Table 4-6. 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.

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 Figure 4-5, until the LAR fitting curve of the clustering was 28.12 seconds. The specific time cost of the k-means clustering and the LAR fitting was 18.45 and 5.47 seconds, respectively. The impact of this methodology is the demonstration that a geo-referenced 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.

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

Table 4-6: This table shows the result of the clustering compared with the seeding locations.

4.4 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.

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, LAI, leaf angle, number of leaves or biomass.

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. doi: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. doi: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. doi:10.3390/s130302830

Butkiewicz, T., 2014. Low-cost Coastal Mapping using Kinect v2 Time-of-Flight Cameras, in: Oceans - St. John's, 2014. IEEE, St. John's, NL, Canada. doi:10.1109/OCEANS.2014.7003084

Dhondt, S., Wuyts, N., Inzé, D., 2013. Cell to whole-plant phenotyping: the best is yet to come. Trends Plant Sci. 18, 428–39. doi: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. doi: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. doi: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. doi: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. doi: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: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. doi: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., 2015. 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. doi: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: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.

Reiser, D., Garrido-Izard, M., Vázquez-Arellano, M., Paraforos, D.S., Griepentrog, H.W., 2015. Crop row detection in maize for developing navigation algorithms under changing plant growth stages, in: Robot 2015. Second Iberian Robotics Conference. Lisbon, Portugal, pp. 371–382.

Reiser, D., Vázquez-Arellano, M., Garrido Izard, M., 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.

Reiser, D., Vázquez-Arellano, M., Garrido-Izard, M., Paraforos, D.S., Sharipov, G., Griepentrog, H.W., 2017. Clustering of Laser Scanner Perception Points of Maize Plants, in: Advances in Animal Biosciences: Precision Agriculture (ECPA). pp. 204–209. doi:10.1017/S204047001700111X

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, 1–1. doi:10.1109/TMECH.2017.2663436

Rusu, R.B., Cousins, S., 2011. 3D is here: point cloud library. IEEE Int. Conf. Robot. Autom. 1 – 4. doi: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 16, 618. doi: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. doi:10.1016/j.robot.2011.02.011

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

Yin, X., Noguchi, N., Choi, J., 2013. Development of a target recognition and following system for a field robot. Comput. Electron. Agric. 98, 17–24. doi:10.1016/j.compag.2013.07.005

5 Part IV: Determination of stem position and height of reconstructed maize plants using a time-of-flight camera

Abstract

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

Keywords: 3-D sensors; Kinect v2; crop characterization; agricultural robotics; precision farming; plant phenotyping

5.1 Introduction

One of the most appealing aspects of reconstructing the geometry of an agricultural environment is to obtain information about the crop status without the troublesome manual measurement. Doing so with an efficient investment in resources, such as fuel and working time (Steckel, 2018), would trigger the interest of farmers in this technology. The information provided by the scanned and digitized data would be very useful for decision-making throughout the cropping cycle; considering that it involves precision agriculture practices (Gebbers and

Adamchuk, 2010). Tasks such as soil, crop, weed and yield sensing and mapping are suitable for 3-D imaging systems. Other applications such as agricultural robotics and plant phenotyping for breeding purposes are among the most appealing (Blackmore et al., 2006). However, since research using 3-D imaging systems in agriculture was previously limited, particularly with the once expensive TOF cameras, there is still the need of new methodologies for extracting useful information out of the 3-D data for agricultural applications. Information such as stem diameter, plant height, leaf angle, LAI, number of leaves, biomass, etc. are of particular interest. If the cost of obtaining such information becomes economically accessible, new applications and solutions will come as a result (Vázquez-Arellano et al., 2016a). An off-the-shelf TOF camera such as the Kinect v2 (Microsoft, Redmond, WA, USA) offers a good cost/performance ratio solution for the development of ground-based 3-D imaging system for proximal sensing (Vázquez-Arellano et al., 2016b). In 3-D imaging a point cloud is a set of data points in space, where each point $P(x, y, z)$ is a function of the spatial position (*x*, *y*, *z*) in a Cartesian coordinate system (Rusu and Cousins, 2011).

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

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

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

5.2 Materials and Methods

5.2.1 Hardware and sensors

The 3-D data used in this research was obtained using a robotic platform, developed at the University of Hohenheim, controlled by a joystick that navigated between maize plants in a greenhouse. The TOF camera was mounted at the front of the vehicle at a height of 0.94 m with a downwards angle of 45 degrees. The SPS930 robotic total station (Trimble Navigation Limited, Sunnyvale, USA) tracked the position of the robot by aiming at the Trimble MT900 Machine Target Prism. An Inertial Measurement Unit (IMU) (VectorNav, Dallas, USA) VN-100 was embedded inside the robotic platform and used to measure its orientation while driving.

The 3-D imaging acquisition system is depicted in Figure 5-1 and the technical characteristics of the robotic platform are described in detail by Reiser et al. (2015).
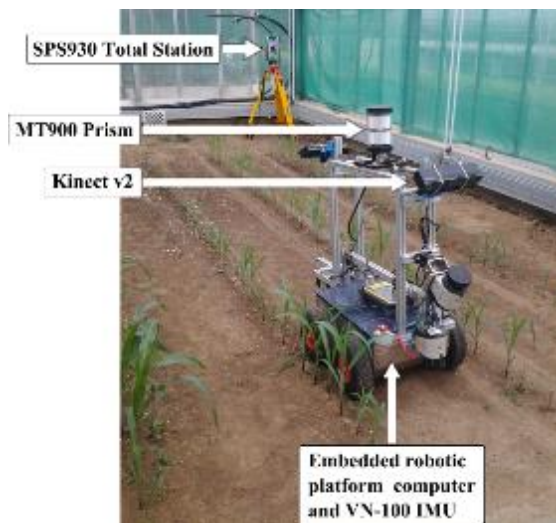
Figure 5-1:   3-D imaging acquisition system with the TALOS robotic platform. The different components for 3-D image, orientation and position data acquisition are marked with arrows and annotations.

5.2.2   Experimental setup

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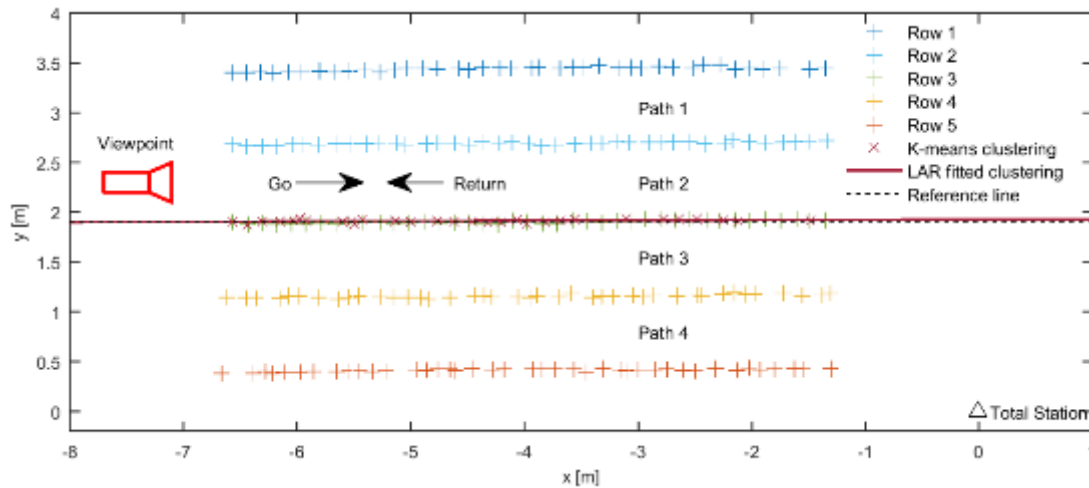
Figure 5-2:   Depiction of the greenhouse maize seedlings (+) and the k-means clustering (x)
             with the least absolute residuals (LAR) representative fitting and reference lines. The
             viewpoint is depicted as a camera plot used to avoid confusion between left and right
       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

5.2.3   Data processing

The raw data for this research is based on the maize plant registration and stitching from a
previous research (Vazquez-Arellano et al. 2018). These point clouds were processed mainly
using the Computer Vision System Toolbox^{TM} of MATLAB R2016b (MathWorks, Natick,
MA, USA). Also, some functionalities of CloudCompare (EDF R&D, 2011) were used for
point cloud processing. In this research, only row 2, 3 and 4 are analysed; row 1 and 5 were
discarded since they were scanned just from one side, in Figure 5-2 it can be seen that they are
both near the edges of the greenhouse where the robotic platform did not fit.

5.2.3.1 Crop row alignment

In the research by (Vázquez-Arellano et al., 2018), an approximation of the plant stems was
done using the *k*-means clustering at a thin layer of 5 mm at the bottom of the plants. The main
intention was to evaluate the accuracy of the reconstruction by relating the clusters with the
seedlings positions (ground truth), but also to know the alignment of the reconstructed crop
rows by fitting a first order polynomial curve considering the least absolute residuals (LAR) to
minimize the effect of extreme values on the fit. This line provides a representative angle of the
crop rows relative to the *x* axis of the total station coordinate system. During the data acquisi-
tion, the *x* axis of the total station coordinate system was intentionally aligned as accurate as
possible with the crop rows, in order to facilitate the data processing. However, in open field

conditions the crop rows are rarely aligned with coordinate system of the geo-referencing system. Therefore, the LAR fitting line was used to transform the reconstructed point cloud by rotating it until it is aligned with the x axis. In Figure 5-2 it can be seen that the angular difference between the LAR fitting of row 3 and the reference line (parallel to the x axis) is minimal (0.0046 degrees). The row 3 clustering was used as a representation of the whole maize rows since it is in the middle row. Also, in Figure 2, the camera plot (in colour red) represents the viewpoint to avoid confusion regarding the left and right side of the crop row. Since there are 3-D reconstructions while going and returning, the left and right side would be different (if a viewpoint was not established) depending on the driving direction.

5.2.3.2 Plant stem position estimation methodology

The methodology for stem position estimation proposed in this research was based on reconstructed crop rows using the methodology proposed by Vázquez-Arellano et al. (2018). The generated point clouds were used to obtain data point density histograms using a bivariate approach by pairing the x and y values of every point in the point cloud (see Figure 5-3b). The registration and alignment of the point clouds was done pairwise, that means with two crop row point clouds at a time (generated from different perspective views) until all of them (4 in total) were aligned, merged and filtered (see Figure 5-3a).
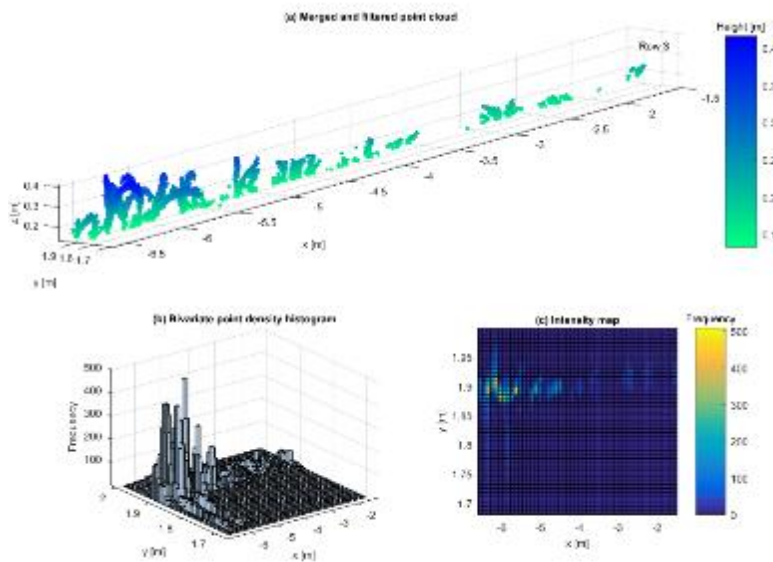

Figure 5-3:   Result of the registration and alignment of four point clouds after (a) merging and filtering. The bivariate (b) point density histogram with another representation as an (c) intensity image, where the warmer squares indicate a local maxima that is related with the plant stems.

Previous research, such as the one by Lu et al. (2015), used univariate point density histograms for detecting the stem of a single plant from different 3-D perspective views. They considered the stem of the plant as the mean value of the histogram plus 3 times the standard deviation. This approach of using univariate point density histograms works well for single plant stem segmentation but it is not suitable for multiple plant analysis. The main reason is that with the univariate approach in a maize row, the information about the position of the individual stems in one axis (x axis in Figure 5-4) is provided as a form of a local peak. But in the perpendicular axis (y axis in Figure 5-4) that information about the position of the individual stems would not be provided, because the y axis value would be the same for all plants since the histogram has a single peak value for the whole maize row. In order to know the position of the individual stems, a unique x and y value must be provided.
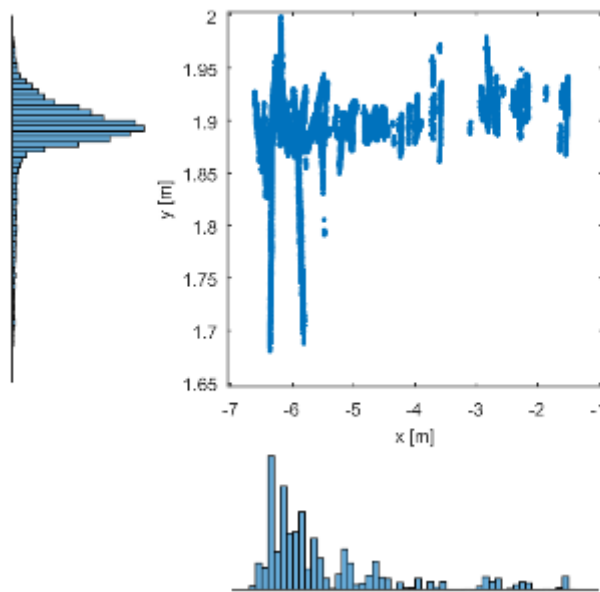
Figure 5-4:   Univariate point density histogram in the x and y axis of the point cloud depicted
in Figure 5-3a.

In single plant morphological analysis, a precise plant stem estimation is an early step in the pipeline to extract other plant parameters, however, in high throughput morphological analysis it is often avoided, rather estimating overall parameters such as biomass or height profiles, due to the difficulty of its detection. This difficulty relies on the fact that plant stem estimation is not an aim by itself, rather an objective to reach the main aim which is single plant segmentation. Single plant segmentation requires complex 3-D imaging algorithms (Reiser et al., 2018) or even machine learning. In this research, a precise plant stem estimation was considered as a

prerequisite for further plant morphological analysis. For instance, in the case of maize, it is difficult to automatically obtain the stem diameter, segment the leaves or calculate the height if the stem position was not previously determined. The methodology for stem position estimation proposed in this research is shown in the flowchart of Figure 5-5.
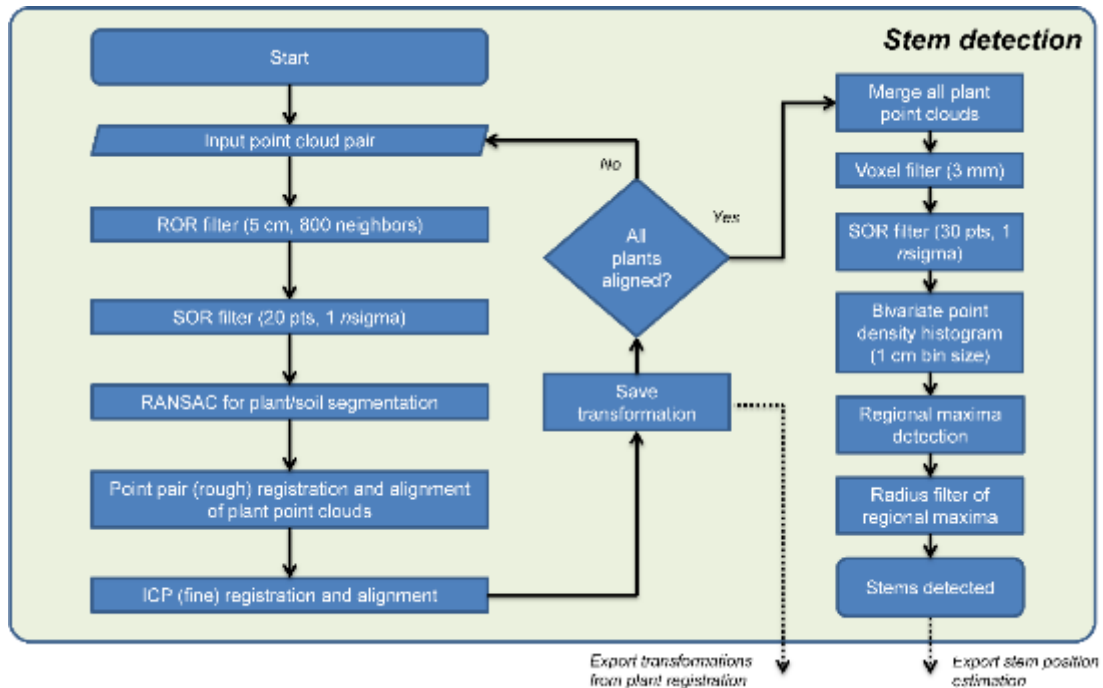
Figure 5-5: Stem detection flowchart showing the pipeline process starting with the input point cloud pair, through the filtering and merging, until the stem detection based on bivariate point density histograms.

The process started by importing a point cloud pair (a) of two rows, later, filtering each point cloud with a radius outlier removal (ROR) filter and statistical outlier (SOR) filter– their input parameters are shown in Figure 5-5. Then, computing the RANSAC algorithm (Fischler and Bolles, 1981), for each point cloud pair, with the maximum distance from an inlier to the plane set to 0.5 times the theoretical distance between the plants (65 mm). Next, aligning them with a point pair (rough) manual registration, and finally using the Iterative Closest Point (ICP) (Besl and McKay, 1992) for (fine) registration and alignment. This process was performed three times, in order to align the four point clouds and merge them together.

The disadvantage of merging several point clouds was that the errors (e.g. due to the point cloud assembly, TOF camera accuracy, moving leaves) accumulate. It was noticed that after merging 4 point clouds, some of the plants appeared thicker or slightly blurred as they were in reality. Therefore, an assessment of the accuracy of the point cloud alignment was necessary to provide

an idea of how well they overlapped. For this, the Cloud-to-Cloud (C2C) method (Girardeau-Montaut et al., 2005) was used to measure the quality of the overlapping by measuring the average distance of one point cloud compared with another one set as a reference, it is also used for detecting changes, therefore this method can be useful for plant growth monitoring. For every point in the compared point cloud, the closest point was defined in the reference point cloud, then the absolute distance was computed and the mean and standard deviation of all the point distances was calculated. CloudCompare software was used to compute the distances and to calculate the mean and the standard deviation. The greater the distance error, the more different the point clouds were. However, it was not possible to distinguish the source of the difference, which could be due to errors in the point cloud generation, moving plants, 3-D perspective view and other parameters.

After the aligned point clouds were merged, a subsequent subsampling was applied using a voxel grid (3 mm × 3 mm × 3 mm) filter, and then, a SOR (20 pts. 1 *n*sigma) filter for noise removal, where *n*sigma means *n* times the standard deviation. At this point, most of the plants point clouds were a continuum without flying (veil) points (Steder and Konolige, 2011) orbiting them. However, the main disadvantage of applying noise reduction filters was that the ToF camera was able, in some cases, to obtain 3-D data of small plants, but since the density and cohesion of those points was low, they were removed by the filter. After filtering, the remaining plant points were used to create a point density histogram with a bin size of 1 cm^2. The resulting bivariate point density histogram was plotted using the shape-preserving Piecewise Cubic Hermite Interpolating Polynomial (PCHIP) method (Kahaner et al., 1989) for interpolation.

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

To remove the regional maxima outliers, a filtering based on the radius filter was implemented. The regional maxima with the largest point count, in other words: the tallest peak, was considered as the starting reference point, since it was the one with the most probabilities of being a stem. The next stem position was then estimated in one direction of the *x* axis, by adding 130 mm, and the maximum regional maxima searching distance was stablished with Equation 5-1.

$$max\_dist = spacing\_factor * plant\_spacing$$
(Equation 5-1)

In this research the maximum distance was 13 mm $* 3 = 39$ mm, therefore, the spacing factor
was 3 and the plant spacing 13 mm. The regional maxima points lying within the maximum
distance (*max_dist*) radius were sorted by the shortest Euclidian distance to the expected stem
position. Then, the closest points were filtered out if their count was bigger than a threshold (set
to 20 points). After all the regional maxima points were analysed in one direction of the *x* axis,
the reference point was once again returned to the regional maxima with the largest point count,
in order to perform the same procedure but in the opposite direction. The disadvantage of merg-
ing several point clouds was that the roughness of the plant surface and edges increases, due
the differences between them, thus increasing the count variability within neighbouring bins
resulting in more detected (noisy) peaks. The main part of the functionality of the filter is better
described with Algorithm 5-1, where *pi.X* and *pi.Y* are related with the bivariate approach pre-
viously mentioned.

Algorithm 5-1 Radius Filter

Require: *pointCloudPlants, nextPlantPosition, maxDist*

for all *points p in pointCloudPlants* **do**

    dx = p_i.X − nextPlantPosition.X

    dy = p_i.Y − nextPlantPosition.Y

    point_dist = euclideanDistance (dx, dy)

    if *point_dist < maxDist* **then**

        pointCloudRadiusInliers (p_i) = PointCloudPlants (p_i)

    end if

end for

5.2.3.3 Plant height estimation methodology

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

This approach applies when the maize plants do not overlap each other (V1-V2 growth stage), however, when the leaves are already long enough to invade the space of the neighbouring plants, the previously mentioned approach do not apply any more. Therefore, a flexible new approach was considered for plant height estimation and it is better described in Figure 5-6.
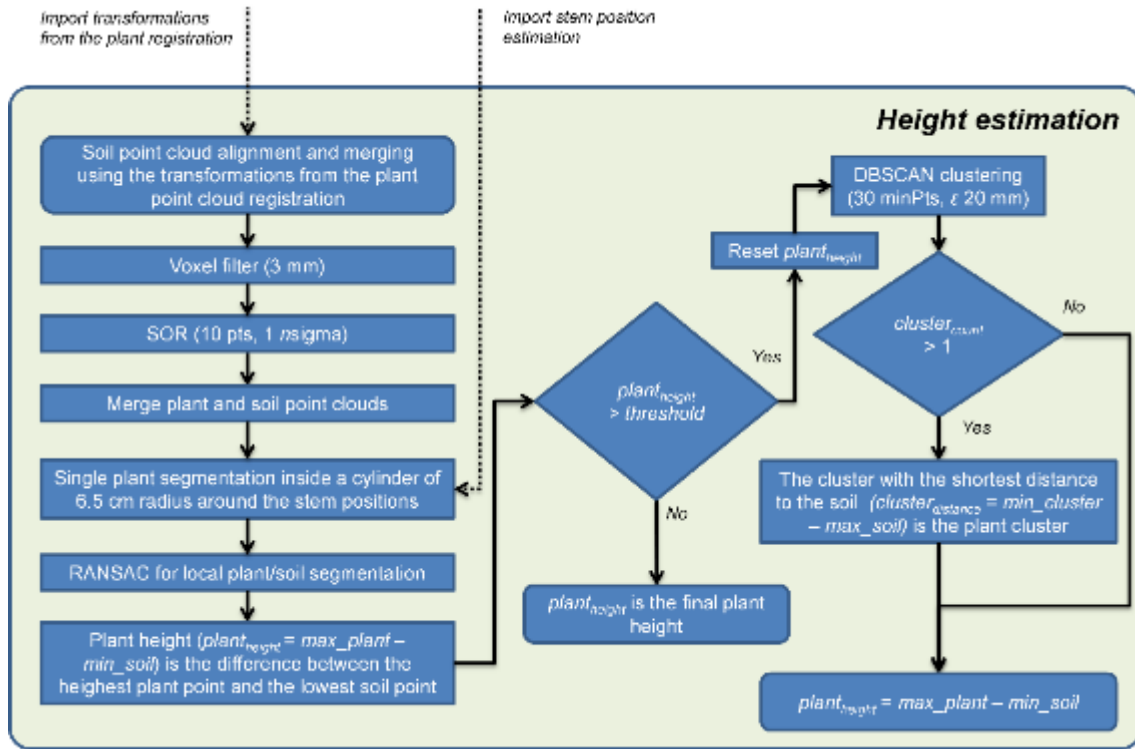
Figure 5-6:   Plant height estimation flowchart showing the pipeline process starting with the soil point cloud alignment, using the imported transformations for plant registration, single plant segmentation, plant clustering and height estimation.

The methodology for plant height estimation started by importing the rigid transformations used to register the plant point clouds in the stem estimation process. Those same transformations were necessary to align and merge the soil point clouds. It is true that the soil point clouds were not necessary for the stem estimation process, however, they were indispensable for the plant height estimation since they provided the reference point for calculating the distance to the point with the maximum height. Additionally, it is important to emphasise that when the plants were initially segmented, using the RANSAC algorithm, in the plant stem estimation pipeline, some plant point were lost, as stated by Garrido et al. (2015), and classified as soil points (which was preferable as soil points classified as plant points), particularly the near soil surface stem points. Those points were recovered here in the plant and soil merging process. After the merging, it was required to perform a filtering, due to duplicate and flying points, using a voxel grid (3 mm $\times$ 3 mm $\times$ 3 mm) and SOR (10 pts. 1 $n$sigma) filters, respectively. At this point, with the cleaned

point cloud from the final merging, the single plant segmentation process could be performed by importing the estimated plant stem positions.

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

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

In order to solve the problem of having the segmented target plant together with an invading leaf or leaves from the neighbouring plants, a clustering algorithm was required to avoid using the invading leaves for the local plant height calculation, instead, the highest z value of the plant cluster plant must be obtained. Algorithms, such as the k-means, that require to specify a priori the number of clusters in the data subset would have limited success, since the number of points and location of the foreign leaf or leaves in the cylinder space could make the previously mentioned algorithm to fail, since it is difficult to predict the number of invading leaves and their location. The DBSCAN was selected because of its flexibility with the clustering of arbitrary shapes and robustness with noisy point clouds. The DBSCAN algorithm requires, as input parameters, the minimum number of points (*minPts*) in the neighbourhood (set to 30 points) and the maximum distance between points (ε) in a cluster (set to 20 mm). The cluster with the shortest distance, from the lowest point of the cluster (*min_cluster*) to the highest point of the soil (*max_soil*), was considered as the plant cluster. Finally, the plant height was calculated with the difference between the maximum height of the plant cluster (*max_plant*) and the minimum height of the soil point cloud (*min_soil*).

5.2.3.4 Plant height profile estimation methodology

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

1. Rasterize the plant point clouds using a 390 mm grid, three times the inter-row plant separation of 130 mm, projecting the maximum height of every cell in the z direction.

2. Rasterize the soil point cloud using a 3 mm grid, thus maintaining the same point density, and projecting the minimum height of every cell in the z direction.

3. Generate a mesh, or 2-D soil profile model, from the soil point cloud using a thinner grid of 50 mm to reduce the roughness and sensibility to noise.

4. Compute the cloud to mesh distance (nearest neighbour) of every rasterized plant point to the meshed soil.

5.3   Results and Discussion

5.3.1   Plant stem position estimation

The merging of the four point clouds from different 3-D perspective views provided, in most of the cases, complementary information for the construction of a better plant representation as seen in Figure 5-7. It can be seen the difference between the point cloud pairs, where the leaf segments of some plants were missing in one point cloud, while visible in the other. Moreover, there were entire plants that were missing in one point cloud, while visible in the other. For example, in Figure 5-7a, it can be seen that the last plants at row 4 ($x = $ -1.5 m, $y = $ 1.25 m) were not visible in the scan from the left side and barely visible for the right side while going, meanwhile, in Figure 5-7b, the same plants were very well defined while returning from both sides. Also, in Figure 5-7a, in row 2 ($x = $ -1.4 m, $y = $ 2.7 m) the last plant was visible in the scan from the right side, but not in the one from the left side.
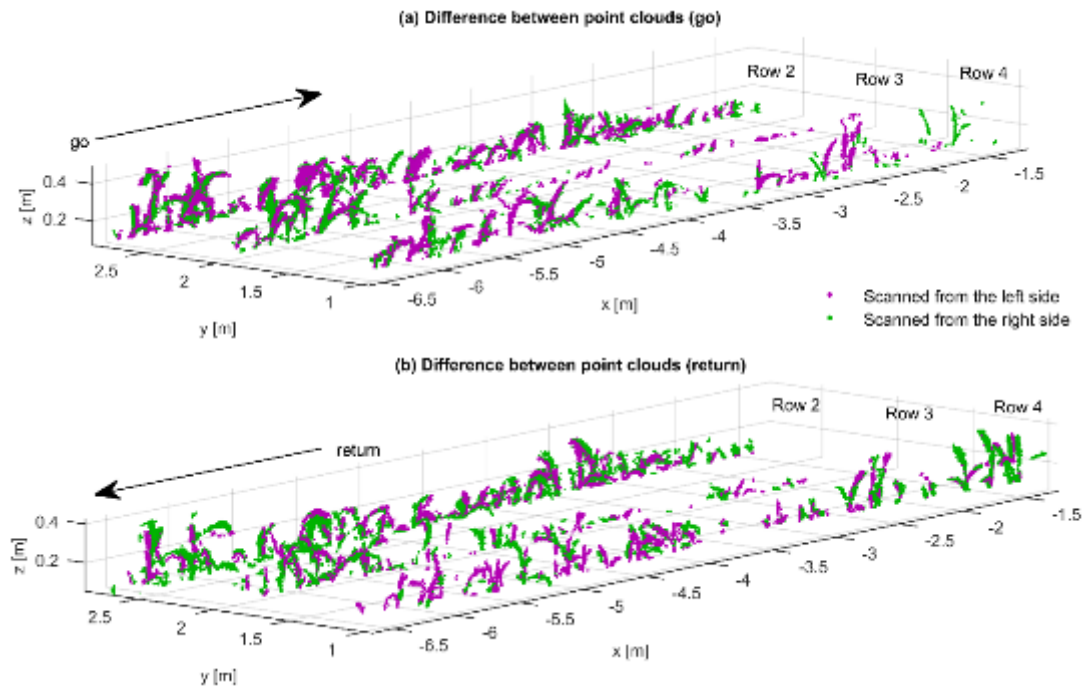
Figure 5-7:   Registration and alignment of point cloud pairs after the ICP algorithm. The point clouds in magenta were reconstructed when the robotic platform drove through the path at the left side of each crop row, while the ones in green were reconstructed when it drove at the right side. The driving direction was (a) go and (b) return.

The merging of several point clouds was necessary due the inherent occlusion in maize crop rows since some parts were out of the reach of the perspective view of the TOF camera, but as previously mentioned, the disadvantage was that miscellaneous errors made the plant thicker as it is in reality. Table 5-1 shows the quantification of the overlapping differences, where the worse one occurred at row 4 with a mean distance error of 30 mm while the best was 11 mm in row 3. However, the average mean and standard deviation of 19 mm and 21 mm respectively, represent a good quality in the overlapping. It must be clarified that the benefit of a better representation of the crop row (in the form of a merged point cloud) outweighs the cost of a small reduction in overlapping quality when the individual point clouds were generated with high precision (as it is the case in this research), but if the merged point cloud lacks precision, such as drifting or wrong registration, the cost-benefit would be low. Also, if the overlapping percentage between subsequent 3-D images during the registration and stitching process is high, more information is obtained and thus less crop row point clouds would be required. In this research the overlapping was not high, particularly at the end of the rows, therefore, four crop row point clouds were required for a better representation.

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

Table 5-1: C2C results of the point cloud overlapping. It is shaded in colour gray the merged point cloud that was used for stem detection and plant height estimation

The interpolated point density histogram shown in Figure 5-8a was the result of the computation of the regional maxima of the bivariate point density histogram. The number of detected regional maxima points was very sensitive, aside from the bin size, to the roughness of the point cloud. In other words, if the point cloud that represents a leaf was either noisy, blurred or incomplete, it would produce more regional maxima points since the local bin count is more uneven. It was desirable to have less regional maxima points since it increases the accuracy of the radius filtering. In the case of the bin size, in this research a square bin size of 1 cm^2 was used, but if a bigger bin size was used, it would have facilitated the stem detection, due to a reduced number of regional maxima, but at the price of a degradation in precision.

Since there were more regional maxima than number of stems, a radius filter was implemented to remove the outliers. After filtering, most the regional maxima remained at the peaks, while the low-lying points (most of them outliers) at the foothills of the high peaks were removed. The result after applying the radius filter is shown in Figure 5-8b.
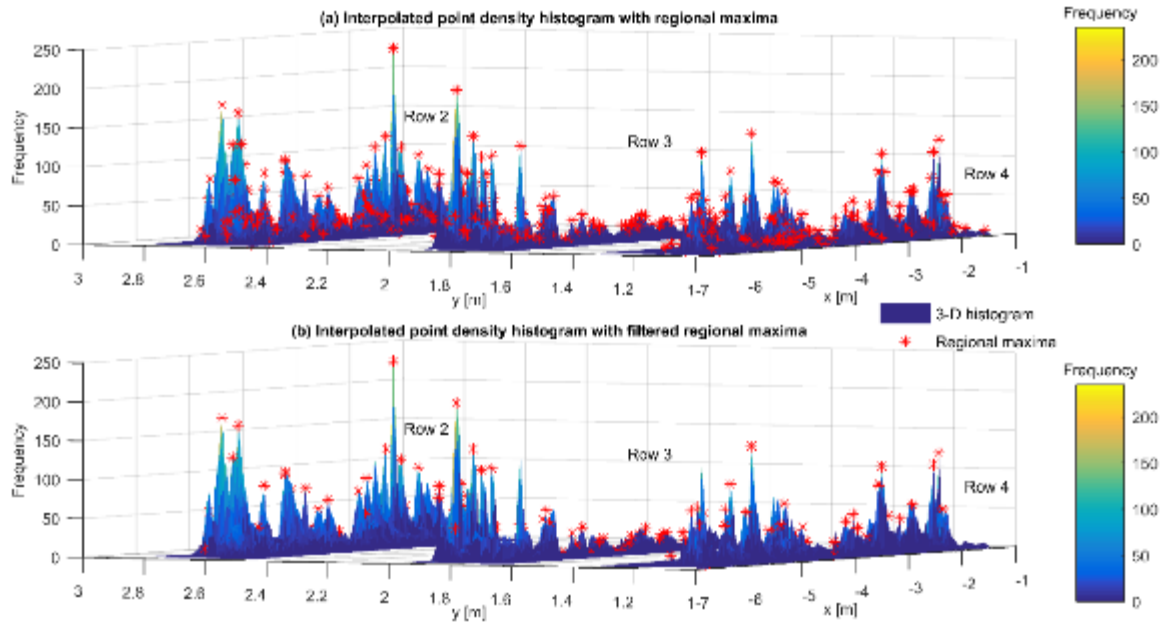
Figure 5-8:   Interpolated point density histogram with regional maxima (*) (a) before and (b)
after the radius filtering.

As it can be noticed in Figure 5-9, the mean error closely correlated with the standard deviation
of the seeding; the more unevenly the seeding was, the larger the stem position estimation error,
therefore, it is not coincidence that the row with the worse standard deviation during seeding
(row 4) had the worst mean error (34 mm).
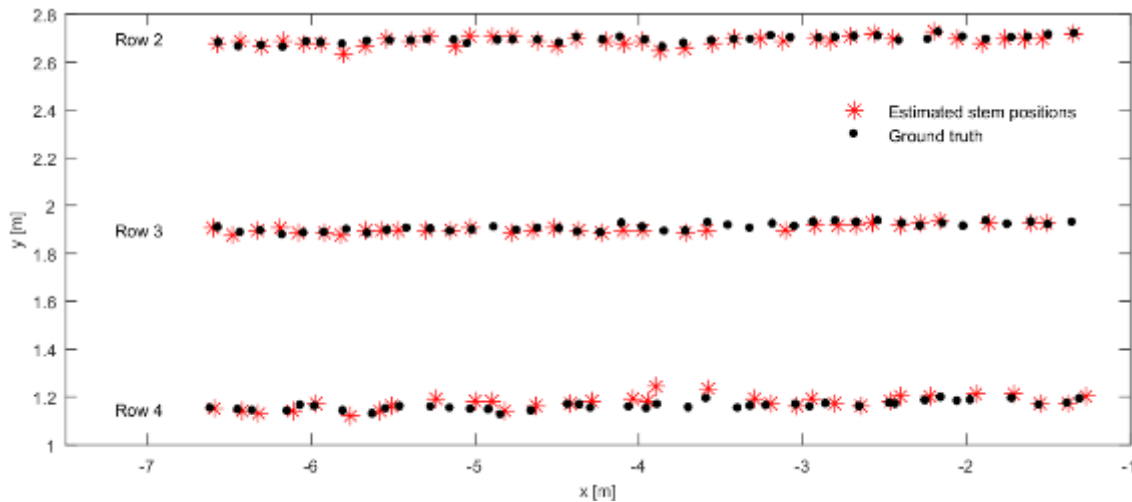
Figure 5-9:   Plant stem position estimation assessment. The ground truth are the seedling
positions measured with the robotic total station, with sub-centimetre accuracy, after
the seeding.

The radius filter algorithm made an initial estimation, to detect a new plant position, expecting
to be located 130 mm apart from the reference plant in the *x* direction. Therefore, the theoretical

distance between plants is used as an input for the radius filter. The plant stem position was estimated with an average mean error and standard deviation and RMSE equal to 27 mm, 14 mm and 34 mm respectively (see Table 5-2). The worst mean error, as expected, was in row 4 due to the high standard deviation in the seeding, which is visually evident in Figure 5-8 where the ground truth plot of row 4 is more scattered compared to row 3. Also, the high standard deviation in the seeding of row 4 put the algorithm into more difficulties, as seen in the less precision of the estimated stem positions compared to row 3 (see Figure 5-9). Additionally, in row 3 there were 8 false negatives that were eliminated either by the RANSAC or by the filtering algorithms. Without considering the false negatives, the algorithm was able to correctly detect approximately 95% of all the stem positions.

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

Table 5-2: Plant stem position estimation.

This process was done iterative until all plant positions had been processed. Figure 5-10 shows the distance error histogram with the 95th percentile equal to 4.8 mm. The distance error was best fitted by a Rayleigh distribution with a scale parameter of 21 mm. It was noticed that with the 3-D imaging system used in the research, it was difficult to detect plants with heights lower than 130 mm or with stem diameter lower than 1.5 mm.
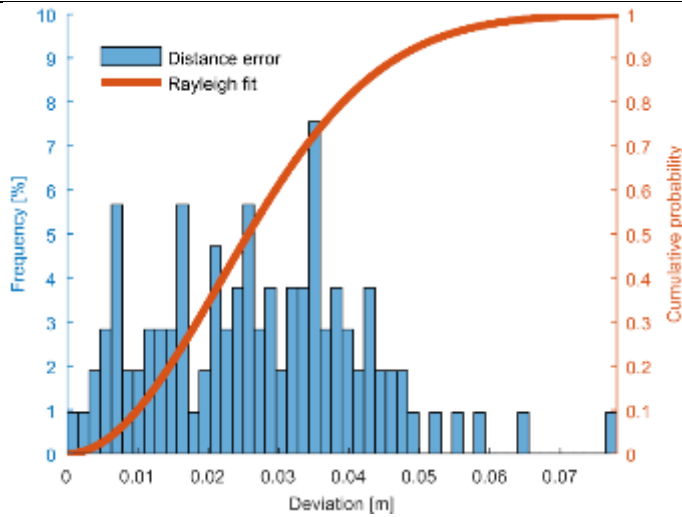
Figure 5-10:   Distance error histogram showing the deviation of the estimated heights in all the analysed maize rows. The cumulative probability was best fitted by a Rayleight distribution.

5.3.2   Plant height estimation

After the stem position was estimated, as seen in Figure 5-11a, the next step was to use this information for individual plant segmentation. In Figure 5-11b, the 65 mm radius cylinders are depicted on the final merged point cloud (four plant point clouds merging together with four soil point clouds merging) that were used as boundaries for the individual plant segmentation. Some plant points of long leaves fall outside the cylinder as seen in Figure 5-11b (two long leaves between $x = -5.5$ m and $x = -6.5$ m), in this cases, the plant height would be underestimated because the real maximum height lies on the outlier. Therefore, this approach could start with considerable inaccuracies with plants from V4 growth stage onwards, since a considerable part of the leaves would lay out of the cylinder. However, as previously said, since plants of the datasets presented in this research are mostly between V1 and V3 growth stages, the overall height estimation methodology was not affected by those few underestimations.
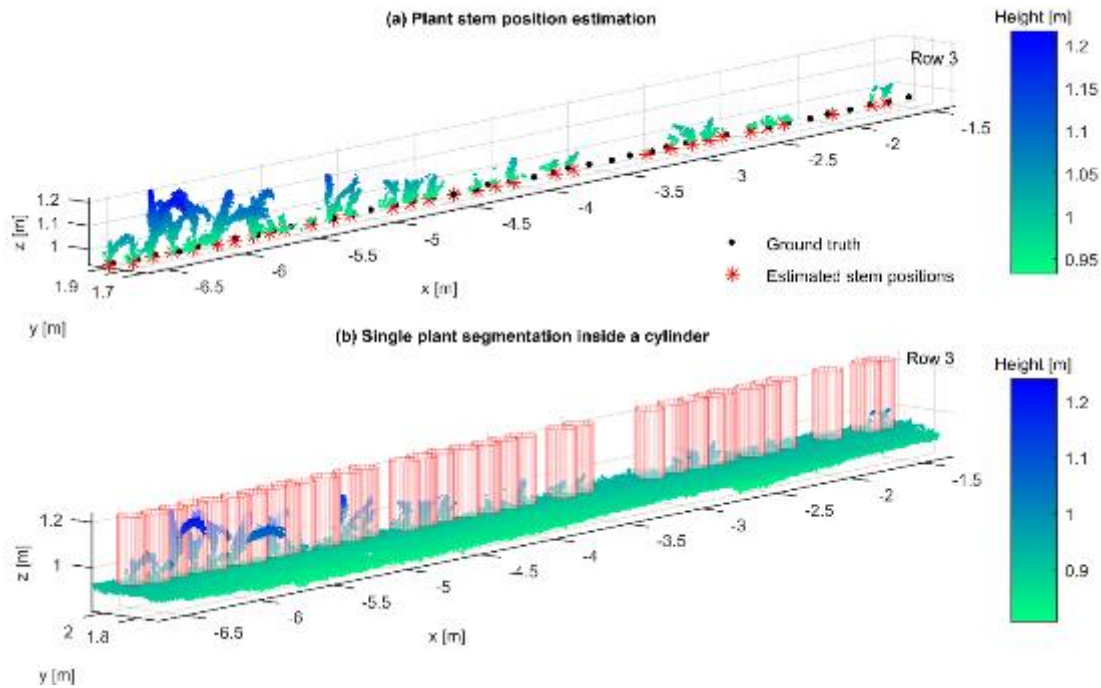
Figure 5-11:   Point cloud with (a) plant stem position estimation compared with the ground
truth. After the stem positions were estimated, they were used to project cylinders
around them (b), for single plant segmentation. Every plant point outside its cylinder,
was consider an outlier. The *z* coordinate is related with the real plant height.

After the plants were segmented, a RANSAC algorithm was implemented to separate the plant
points from the soil points in order to discriminate, within the plant points, between the ones
that were part of the local plant and the ones belonging to neighbouring leaves invading the
local plant space. An important part of the proposed methodology was the implementation of
the DBSCAN clustering, without it, large height underestimations would decrease the accuracy
of the height estimation and therefore the viability of the proposed methodology. In Figure 5-
12a it can be seen a single plant clustering, which applied to most of the plants, because as
previously said: most of the plants were inside the boundaries of the cylinder. Figure 5-12b,
Figure 5-12c and Figure 5-12d show the successful discrimination between the target plant and
the invading leaves from the neighbouring plants. In Figure 5-12c, it is shown that Plant 3
(height = 200 mm, stem diameter = 3.5 mm) is between two bigger plants that have occluded it
during the 3-D data acquisition, therefore, the points in the plant point cloud were sparse, how-
ever, the DBSCAN algorithm was able to cluster them together and differentiate them from the
incoming leaf from the neighbour plant. Although, the DBSCAN algorithm was very effective,
it could not correctly cluster plants with high leaf overlapping, additionally, there were several
cases of folded leaves, that were too thin to be detected by the TOF camera, but their end points

(apex or tip) were the highest point of the plant and were taken as a reference for the manual
height measurement.
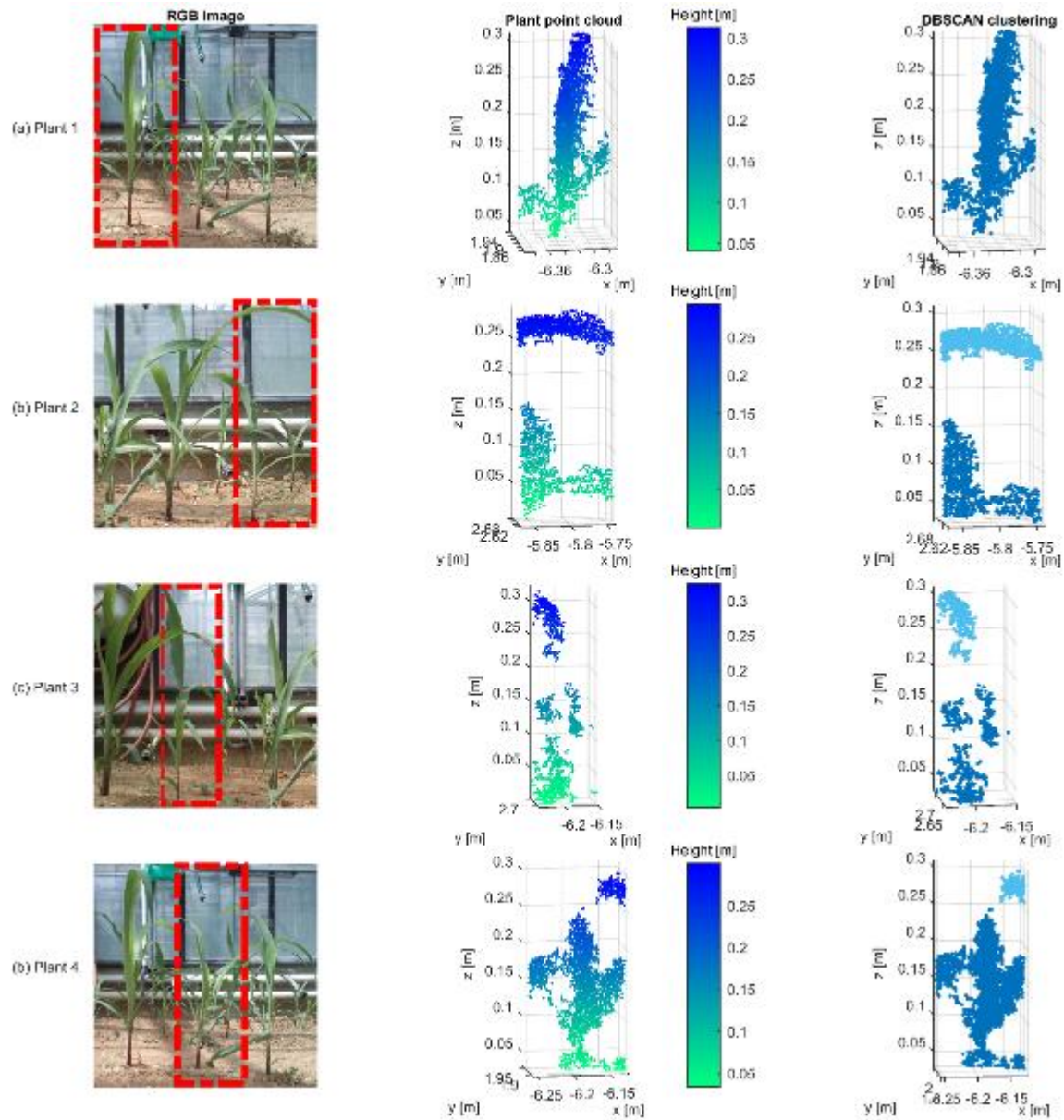
Figure 5-12:   DBSCAN clustering algorithm in 4 different plants (a-d). The dark blue clusters
represent the plant and the bright blue the invading leaf.

To validate the plant height profile the mean height of every row was analysed. Table 5-3 shows
that the average mean error and standard deviation and RMSE were 30 mm, 35 mm and 50 mm,
respectively. The worst mean error was in row 4, which partly was affected by the worsening
of the stem estimation due to the high variability in the plant positions. Aside from that, other
sources of errors were the lack of 3-D data on small plants and folded leaves that produced
underestimated heights. Figure 5-13 shows the distance error histogram with the 95th percentile
equal 0.1 m and its best fit (Nakagami distribution), with the shape and scale parameters equals

to 0.53 and, 0.0026 respectively. The main source of inaccuracies for the height estimation were mainly the lack of 3-D data of the plant, therefore, although the plants higher than 130 mm were detected, the ones smaller than 200 mm or with stem diameter smaller than 30 mm, were more an abstract plant point cloud with not a clear and objective 3-D plant representation. Therefore, neither plant phenotyping nor a very precise height estimation could be performed, signalling the limitations of the 3-D TOF camera.

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

Table 5-3: Plant height estimation.
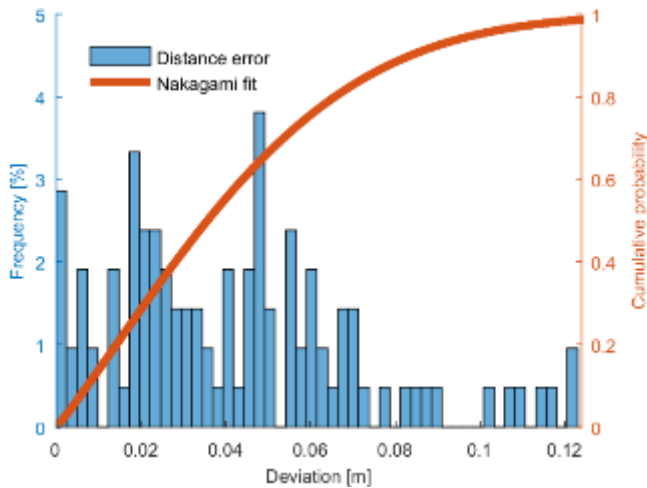
Figure 5-13:   Distance error histogram shows the deviation of the estimated heights of all the analysed plants. The cumulative probability was best fitted by a Nakagami distribution.

5.3.3 Plant height profile estimation

Results of the plant height profile (see Figure 5-14) not just visually correlate with the final reconstructed crop row, but also numerically. The rasterized points of the plant height profile could not be compared with single plant heights, since it was an average of several plant heights in a 0.39 m × 0.39 m cell, therefore, the assessment was done by comparing the mean height of the rasterized points with the mean height of the ground truth per row. Table 5-4 shows very precise results for the height profiles, reaching 6 mm error in row 2 and 8.7 mm in the average of the three rows.

Figure 5-14:   (a) Height profile representation with rasterized plant points and meshed soil point cloud. (b) The merged point cloud closely agrees with the height profile (red line).

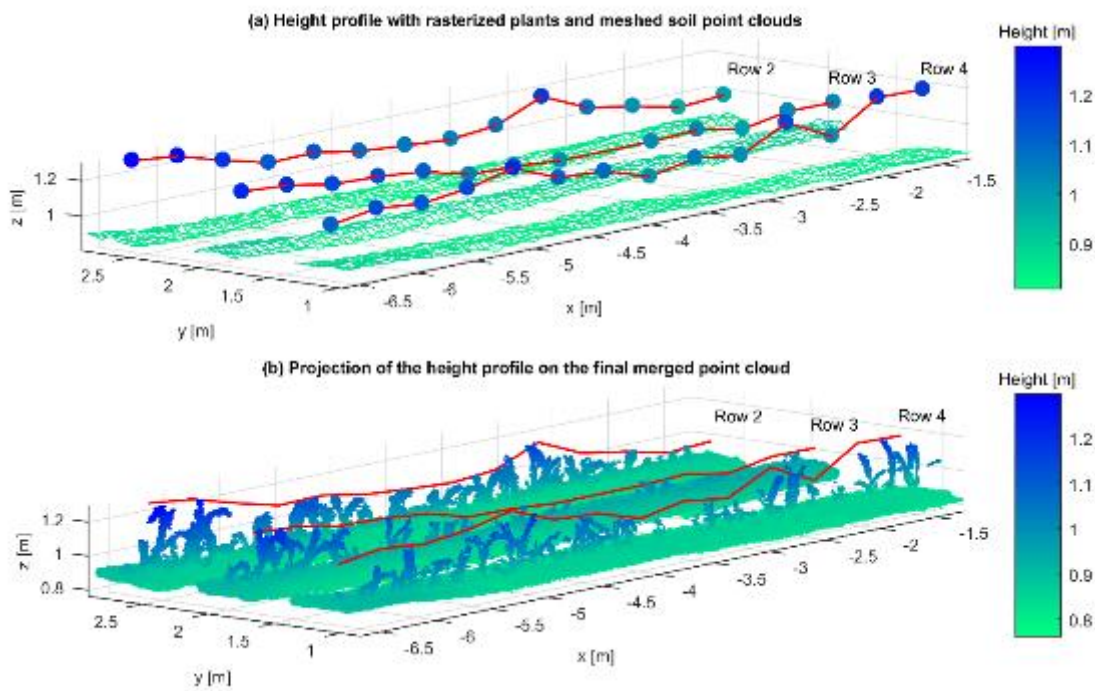| Crop Row	Rasterized plant points [no.]	Maize plants [no.]	Overall mean error [mm]
2	14	41	6
3	13	41	8
4	14	41	12
Average			8.7

Table 5-4: Plant height estimation

5.4 Conclusions

The results demonstrated that the TOF camera-based 3-D imaging system was able to estimate the stem position with accuracies of 27 mm and 14 mm standard deviation. It also provided meaningful information about the plant height profile with an average overall mean error of 8.7 mm. Since the maize plants considered in this research were highly heterogeneous in height, some of them had folded leaves and were planted with standard deviations that emulate the real performance of a seeder; it can be said that the experimental maize setup was a difficult scenario. Therefore, a better performance, for both, plant stem position and height estimation could be expected for a maize field in better conditions since maize heights are more homogeneous and the leaves are unfolded. Ground-based 3-D image acquisition provided important data about the plant stem, and although the acquisition is slow (between 0.02-0.04 m·s^{-1}) compared with an unmanned aerial system, the latter would not be able to obtain plant stem data. Another approach would have been to place the camera in a side-view position in order to obtain more data about the plant stem. Finally, having a 3-D reconstruction of the maize plants using a cost-effective sensor, mounted on a small electric-motor-driven robotic platform, means that the cost (either economic, energetic or time) of generating every point in the point clouds is greatly reduced compared with previous researches. Further research needs to be done with the use of clustering algorithms for single plant segmentation (e.g. Euclidian, min-cut based, region growing segmentation) so that parts of long leaves are not cut off from the plant point cloud, and that plants at a latter growth stage can be also analyzed. Also, non-rigid registration and align-

ment needs to be explored to reduce the effect of the miscellaneous errors that reduce the overlapping between point cloud pairs. Additionally, the LAI with the data generated with this TOF camera needs to be analyzed at different growth stages to assess its correlation since it is a very important plant parameter that needs to be taken into consideration.

Conflict of Interest: The authors declare no conflict of interest.

References

Andújar, D., Dorado, J., Fernández-Quintanilla, C., & Ribeiro, A. (2016). An Approach to the Use of Depth Cameras for Weed Volume Estimation. *Sensors*, *16*(972), 1–11. doi:10.3390/s16070972

Besl, P., & McKay, N. (1992). A Method for Registration of 3-D Shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. doi:10.1109/34.121791

Blackmore, B. S., Griepentrog, H. W., & Fountas, S. (2006). Autonomous Systems for European Agriculture Conference : Automation Technology for Off-Road Equipment. In *Automation technology for off-road equipment*. Bonn, Germany.

Chaivivatrakul, S., Tang, L., Dailey, M. N., & Nakarmi, A. D. (2014). Automatic morphological trait characterization for corn plants via 3D holographic reconstruction. *Computers and Electronics in Agriculture*, *109*, 109–123. doi:10.1016/j.compag.2014.09.005

EDF R&D, T. P. (2011). CloudCompare (version 2.9.1) [GPL software]. http://www.cloudcompare.org/. Accessed 12 May 2017

Ester, M., Kriegel, H.-P., Sander, J., & Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. In E. Simoudis, J. Han, & U. M. Fayyad (Eds.), *Proceedings of the Second International Conference on Knowledge Discovery and Data Minning* (pp. 226–231). AAAI Press.

Fischler, M. a., & Bolles, R. C. (1981). Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, *24*(6), 381–395. doi: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 Sensing*, *7*, 17077–17096. doi:10.3390/rs71215870

Gebbers, R., & Adamchuk, V. I. (2010). Precision agriculture and food security. *Science*, (327), 828–831.

Girardeau-Montaut, D., Roux, M., Marc, R., & Thibault, G. (2005). Change detection on points cloud data acquired with a ground laser scanner. *IAPRSSIS*, *39*(W19).

Hämmerle, M., & Höfle, B. (2016). Direct derivation of maize plant and crop height from low-cost time-of-flight camera measurements. *Plant Methods*, *12*(50), 1–13. doi:10.1186/s13007-016-0150-6

Kahaner, D., Moler, C., & Nash, S. (1989). *Numerical methods and software*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc.

Lu, H., Tang, L., & Whitham, S. A. (2015). Development of an automatic maize seedling phenotyping platform using 3D vision and industrial robot arm. In *ASABE Annual International Meeting* (Vol. 7004). Louisiana, USA.

Lu, H., Tang, L., Whitham, S. A., & Mei, Y. (2017). A Robotic Platform for Corn Seedling Morphological Traits Characterization. *Sensors*, *17*(2082), 1–17. doi:10.3390/s17092082

Nakarmi, A., & Tang, L. (2012). Automatic inter-plant spacing sensing at early growth stages using a 3D vision sensor. *Computers and Electronics in Agriculture*, *82*, 23–31. doi:10.1016/j.compag.2011.12.011

Reiser, D., Garrido-Izard, M., Vázquez-Arellano, M., Paraforos, D. S., & Griepentrog, H. W. (2015). Crop row detection in maize for developing navigation algorithms under changing plant growth stages. In *Robot 2015. Second Iberian Robotics Conference* (pp. 371–382). Lisbon, Portugal.

Reiser, D., Vázquez-Arellano, M., Garrido-Izard, M., Paraforos, D. S., Sharipov, G., & Griepentrog, H. W. (2017). Clustering of Laser Scanner Perception Points of Maize Plants. In *Advances in Animal Biosciences: Precision Agriculture (ECPA)* (Vol. 8, pp. 204–209). doi:10.1017/S204047001700111X

Ribeiro, A., Nuñez, N., Cantuña, K., & Andújar, D. (2017). 3D monitoring of woody crops using an unmanned ground vehicle. In *Advances in Animal Biosciences: Precision Agriculture (ECPA)* (pp. 210–215). Edinburgh, UK. doi:10.1017/S2040470017001200

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

Rusu, R. B., & Cousins, S. (2011). 3D is here: point cloud library. *IEEE International Conference on Robotics and Automation*, 1 – 4. doi:10.1109/ICRA.2011.5980567

Steckel, T. (2018). *Entwicklung einer kontextbasierten Systemarchitektur zur Verbesserung des kooperativen Einsatzes mobiler Arbeitsmaschinen*. Universität Hohenheim.

Vázquez-Arellano, M., Griepentrog, H. W., Reiser, D., & Paraforos, D. S. (2016). 3-D Imaging Systems for Agricultural Applications — A Review. *Sensors*, *16*(5), 618. doi:10.3390/s16050618

Vázquez-Arellano, M., Reiser, D., Garrido, M., & Griepentrog, H. W. (2016). Reconstruction of geo-referenced maize plants using a consumer time-of-flight camera in different agricultural environments. In A. Ruckelshausen, A. Meyer-Aurich, T. Rath, G. Recke, & B. Theuvsen (Eds.), *Intelligente Systeme - Stand der Technik und neue Möglichkeiten* (pp. 213–216). Osnabrück, Germany: Gesellschaft für Informatik e.V. (GI).

Vázquez-Arellano, M., Reiser, D., Paraforos, D. S., Garrido-Izard, M., Burce, M. E. C., & Griepentrog, H. W. (2018). 3-D reconstruction of maize plants using a time-of-flight camera. *Computers and Electronics in Agriculture*, *145*(August 2017), 235–247. doi:10.1016/j.compag.2018.01.002

6   General discussion

6.1   3-D imaging in agricultural applications

An extensive review of 3-D imaging systems for agricultural application was done. Here it was mentioned that 3-D imaging hardware is getting cheaper and software tools for 3-D image processing, that were not previously available, such as ROS, OpenCV, PCL and CloudCompare are under an open source license. The usefulness of 3-D imaging in agriculture relies on the fact that plants have a complex architecture that a 2-D approach often fails to fully describe. Additionally, 3-D sensors are becoming smaller, smarter and cheaper. Therefore, technology breakthroughs are possible if enough research were commercialized since it is already economically justified to use 3-D sensors for producing agricultural products. Applications like precision farming, 3-D phenotyping for plant breeding, collision awareness and driving assistant systems are some examples where 3-D imaging system have a lot of potential. The advantage of 3-D imaging, as mentioned in Part I, is precisely the enhanced perception to describe not only the plant architecture, but also the complex environment where soil, weeds, trees, obstacles, machinery etc. are common elements in the agricultural scenario.

Although the CCD was invented in 1969, at the time of the review paper in Part I, there were very few commercial implementations of 3-D imaging systems in agriculture (Ruckelshausen, 2012). Additionally, those limited applications were economically expensive. In Part I it was mentioned about the potential of the low-cost TOF cameras, like the Kinect v2, for agricultural applications due to the superior technical characteristics related with depth image pixel resolution and price. Also, the depth measurement accuracy of this TOF camera is not very different from the other commercial TOF cameras (Kazmi et al., 2014; Klose et al., 2009). There are some publications even concluding that the Kinect v2 provides superior overall results compared to more expensive TOF cameras (Laukkanen, 2015). It has some disadvantages such as its reduced sensing range (0.4 m to 4.5 m), lack of full robustness in open field environments, synchronization difficulties when building a multiple-camera system (i.e. TOF-TOF) and the lack of a robust casing with good quality enclosure. However, cost-effective devices are needed in agriculture since it is an industry with less economical resources available for high-end technological applications compared with the automotive, military or medical industries.

The plant phenotyping bottleneck can be relieved by the use of 3-D imaging cameras, since some research has already been conducted with low pixel resolution TOF cameras (Alenyà et al., 2014; Kizma et al., 2012; Lu et al., 2015), the plant under consideration should be taller

than 130 mm and with a diameter larger than 1.5 mm, as found in Part IV, if a maize plant is considered. Leafy vegetables were recently successfully reconstructed by Hu et al. (2018), and with it, they were able to monitor the plant growth and also measuring, aside from plant height, total/projected leaf area and volume. Additionally, some research has been done using the structured light cameras for low-cost plant phenotyping (Paulus et al., 2014).

3-D imaging systems can also play an important role in precision farming. In Part III it was shown that entire maize plants rows can be accurately reconstructed, however, it must be clarified that the benefit of a better representation of the crop row (in the form of a merge of two or more point clouds) outweighs the cost of a small reduction in quality, due to overlapping errors, when the individual point clouds are generated with high precision; but if the merged point clouds lack precision, such as drifting or wrong registration, the cost-benefit would be low. Also, if the overlapping percentage between subsequent 3-D images during the registration and stitching process is high, more information is obtained and thus less crop row point clouds would be required for merging. For high throughput phenotyping, it could be necessary to increase the 3-D image sampling frequency to allow more overlapping between successive image pairs, also, some artificial markers, as recommended in Part III, could be hammered into the soil and reconstructed together with the plants. The well-defined shape of the marker could help to recalibrate the registration and stitching process. The reconstructed maize plants are very useful for growth monitoring, also, online applications based on the real-time height profile measurement could be implemented, as well as offline applications based on 3-D maps.

Agricultural robotics has been obtaining attention, not just from academic institutions, but also from well-established original equipment manufacturers (OEMs) and start-ups. In Part I it was mentioned that the number of research publications related with 3-D imaging systems in agriculture has increased rapidly in the last years. Some researches consider TOF cameras a mature technology that is widely being adopted to provide sensory input to robotic applications in agriculture (Alenyà et al., 2014). Also, some OEMs have invested in multi-year research projects related with agricultural robotics. An example is Deepfield Robotics, which relies on a Kinect v2 for plant phenotyping in their Bonirob field robot (Deepfield Robotics, 2016), but for acquiring stable images, they rely on a shadowing device that houses the TOF camera together with an artificial light source.

6.2 Validating generated point clouds based on 3-D image registration and stitching

In Part II it was shown (Figure 3-2) that it was possible to obtain 3-D data using a TOF camera in different agricultural scenarios. One important scenario for 3-D imaging that was so far not yet fully researched is the night environment, where robotic applications could extend the working hours of a farmer to a 24/7 operation, and the reconstruction was shown within this research in Figure 3-3.

The results of Part III validate the accuracy of the generated point clouds, obtained in the daytime, using a low-cost TOF camera for reconstructing agricultural environments, and with it the recognition of structures of interest for precision farming applications. The methodology proposed for maize plant reconstruction was based on the balance between the plant and soil point ratio expressed in Equation 4-6. This balance between plant and soil points in every 3-D image allowed a more accurate registration that stitched a sequence of images without having a considerable drifting effect in the end part of the final point cloud. This drifting effect is well known problem in 3-D image registration and stitching, therefore other researchers developed methodologies to correct the final point cloud affected by the drifting effect (Ribeiro et al., 2017). The methodology proposed in this research avoids this effect by spending more computational resources in the pairwise registration. The maize plant elements are clearly visible in the images presented in Part III, where the leaves are very well defined. It was punctually stated the limits of the TOF camera for plant measurement. Due to the limited proximal range of 0.4 m, small plants cannot be detected by the sensor. For maize plants, it was difficult to detect plants with heights lower than 130 mm or stem diameter lower than 1.5 mm. However, it must be clarified that those limit values were found after filtering the point cloud. In reality, the TOF camera was able to obtain 3-D data of plants as small as 90 mm high and stem diameter of 1.3 mm (see Part III). However, in practice those points were so few and scattered, that were lost either in the plant and soil segmentation process or in the noise filtering. Additionally, it was difficult for the human eye to recognize a plant out of scattered points, therefore, it is not expected for an algorithm to be able to obtain plant parameters (aside from an approximation of the plant height).

During the experiments, the position of the camera was at a height of 0.94 m with a downwards angle of 45 degrees. It was very important to have a well-defined and dense stem reconstruction (for the methodology proposed in Part III) in order to estimate its position and the height of the maize plant. Another option for positioning the sensor could have been with the TOF camera

pointing to the side of the crop row (Nakarmi and Tang, 2012), either left or right. However, due to the inter-row spacing of 0.75 m, the TOF camera would have been very close to the spacing limits, thus being within range of the leaves from the adjacent crop row that could have occluded it. Additionally, this side-view position limits the depth images to the sole purpose of plant parameter measurement (with the previously mentioned adversities), while the pose proposed in this research allows the potential use of the depth information for other purposes such as robot navigation. In conclusion, the main task for 3-D sensors in agriculture is to use their superior perception capabilities to solve problems where other sensors, such as 2-D cameras, fail due to their intrinsic limitations.

6.3   Feature extraction and validation of stem position and plant height

In Part IV it was shown that it is possible to measure, with high precision, the height profile of a crop row, and also to measure the individual plant height with a mean error and standard deviation of 30 mm and 35 mm, respectively. Plant height is a very important parameter since it provides information about the individual plant growth status. Additionally, as the review of Part I revealed, height measurement could be used in other applications such as guidance and trailer fill level. Also, in Part IV, was mentioned the possibility of measuring LAI from the generated crop row point clouds. This plant parameter is also very valuable since it can be correlated with biomass. Biomass measurement is of great interest for yield estimation.

One of the remarkable results obtained from this work, were the accuracies in the maize plant stem estimation with an average mean error was 27 mm. The histogram approach was used in a research by Lu et al. (2017), but with an univariate point density histogram approach. This approach was initially considered for the Part IV (Figure 5-3), but it was noticed that it was difficult to implement for more than one plant since the stem location in the *y* axis was dependent on the precision of the stem location in the *x* axis. The bivariate point density histogram approach was more robust because the *x-y* location was automatically generated with the data point density histogram, and none of the variables depended on the other. After the regional maxima of the whole crop row was computed a radius filtering was implemented. It was noticed that the accuracy of the filtering was strongly dependent on the plant spacing, and thus the crop row seeded with the worse standard deviation had also the least accurate regional maxima filtering. The filtering was based solely on the Euclidean distances between a guessed location where a plant is expected and a number of regional maxima around it, established by the Equation 5-1.

It was noticed that in order to make more robust the regional maxima filtering, a homogeneous plant height along the crop row was imperative. Contrarily, the plant height was very heterogeneous and sometimes with a high variation from one plant and its immediate subsequent. Therefore, it can be assumed that a precise seeding together with a more homogenous plant height within a row, would increase the precision of the plant stem estimation. Better precision could be reached if another single plant clustering methodology was used instead of the proposed cylindrical boundary approach. However, for practicality, the cylinder boundary took less computational resources and was easier to implement with good precision, in plant stem estimations, with the considered dataset. For further development of the methodology, the estimations could have been used as an initial guessing for a more complex clustering algorithm such as the one proposed by Reiser et al. (2017). This type of algorithm, min-cut segmentation (Golovinskiy and Funkhouser, 2009), could segment entire plants even if their leaves are invading the neighboring space. However, it is still a difficult task to segment overlapping leaves, belonging to different plants and that touch each other.

According to the ancient roman philosopher Luciud Annaeus Seneca: "Fortune is of sluggish growth, but ruin is rapid". Agriculture has been a fundamental part of civilization, the technical improvements in the agricultural practice has relieved us as a human species from time that, rather than spending it in food production, we have spent in the development of science and technology that are the sediments of our modern civilization. However, if we would like to avoid the collapse of our planetary civilization due to the numerous challenges that are threatening our food production, we need to embrace change. The use of the best of our technological developments must be put into test in a primary an ancient activity, agriculture, with the sole pursue of keep feeding ourselves in times of human overpopulation, resource scarcity and climate change. In this work, an effort has been done to provide an insight of the possibilities of a technological development that could help to put into practice, as a first step, a more sustainable agriculture.

6.4   Outlook

In this research, the technical capabilities, of a consumer low-cost TOF camera, to reconstruct agricultural structures and measure plant parameters were studied. For this, a review of different 3-D imaging sensors was done to investigate the state-of-the-art in in agricultural applications. In this review, it was found that there were few publications regarding the use of TOF cameras for agricultural applications due to their high economical cost. The use of a low-cost TOF camera for crop row reconstruction had not been investigated for agricultural applications. Since it

was not designed to perform in the difficult agricultural environments, there were many questions regarding its technical capabilities for obtaining reliable 3-D data. Within this work, it was shown that a low-cost consumer TOF camera has enough technical capabilities to obtain 3-D data that can be used for agricultural applications.

3-D imaging in agriculture can help to close the chasm between the virtual and real world since it allows the development of concepts such as virtual reality, rapid prototyping and digital agriculture (Vougioukas et al., 2014) by generating models derived from scanned data (Hackenberg et al., 2014). For example: the mathematical representation of real agricultural structures can be very useful for the design of agricultural robots, since the CAD models could be embedded in their working contextual environment if the appropriate kinematic equations are used. Additionally, agricultural tasks could also be modelled based on real world 3-D data, helping to develop more efficient processes and facilitating the planning of precision agriculture tasks, among others.

References

Alenyà, G., Foix, S., Torras, C., 2014. ToF cameras for active vision in robotics. Sensors Actuators, A Phys. 218, 10–22. doi:10.1016/j.sna.2014.07.014

Deepfield Robotics, 2016. BoniRob [WWW Document]. URL http://www.deepfield-robotics.com/ (accessed 2.1.16).

Golovinskiy, A., Funkhouser, T., 2009. Min-cut based segmentation of point clouds. 2009 IEEE 12th Int. Conf. Comput. Vis. Work. ICCV Work. 2009 39–46. doi:10.1109/ICCVW.2009.5457721

Hackenberg, J., Morhart, C., Sheppard, J., Spiecker, H., Disney, M., 2014. Highly accurate tree models derived from terrestrial laser scan data: A method description. Forests 5, 1069–1105. doi:10.3390/f5051069

Hu, Y., Wang, L., Xiang, L., Wu, Q., Jiang, H., 2018. Automatic non-destructive growth measurement of leafy vegetables based on kinect. Sensors (Switzerland) 18. doi:10.3390/s18030806

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. doi:10.1016/j.isprsjprs.2013.11.012

Kizma, W., Foix, S., Alenyà, 2012. Plant leaf analalysis using Time of Flight camera under sun, shadow and room conditions, in: IEEE International Symposium. Magdeburg, Germany. doi:10.1109/ROSE.2012.6402615

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.

Laukkanen, M., 2015. Performance Evaluation of Time-of-Flight Depth Cameras.

Lu, H., Tang, L., Whitham, S.A., 2015. Development of an automatic maize seedling phenotyping platform using 3D vision and industrial robot arm, in: ASABE Annual International Meeting. Louisiana, USA.

Lu, H., Tang, L., Whitham, S.A., Mei, Y., 2017. A robotic platform for corn seedling morphological traits characterization. Sensors (Switzerland) 17, 1–17. doi:10.3390/s17092082

Nakarmi, A.D., Tang, L., 2012. Automatic inter-plant spacing sensing at early growth stages using a 3D vision sensor. Comput. Electron. Agric. 82, 23–31. doi:10.1016/j.compag.2011.12.011

Paulus, S., Behmann, J., Mahlein, A.K., Plümer, L., Kuhlmann, H., 2014. Low-cost 3D systems: Suitable tools for plant phenotyping. Sensors (Switzerland) 14, 3001–3018. doi:10.3390/s140203001

Reiser, D., Vázquez-Arellano, M., Garrido-Izard, M., Paraforos, D.S., Sharipov, G., Griepentrog, H.W., 2017. Clustering of Laser Scanner Perception Points of Maize Plants, in: Advances in Animal Biosciences: Precision Agriculture (ECPA). pp. 204–209. doi:10.1017/S204047001700111X

Ribeiro, A., Bengochea-Guevara, J.M., Conesa-Muñoz, J., Nuñez, N., Cantuña, K., Andújar, D., 2017. 3D monitoring of woody crops using an unmanned ground vehicle. Adv. Anim. Biosci. 8, 210–215. doi:10.1017/S2040470017001200

Ruckelshausen, A., 2012. Neue Sensorentwicklungen: Ein technischer Blick auf Pflanzen, Sensoren und Daten, in: Sensorik Im Pflanzenbau. Quedlingburg, pp. 1–20.

Vougioukas, S., Arikapudi, R., Slaughter, D., Elkins, R., Ingels, C., 2014. A rapid prototyping design tool for pear harvest-aid platforms utilizing 3D fruit reachability and kinematc modeling. Davis, USA.

Declaration of Authorship

I hereby declare that this doctoral thesis is a result of my own work and that no other than the indicated aids have been used for its completion. All quotations and statements that have been used are indicated. I did not accept the assistance from any commercial agency or consulting firm. Furthermore, I assure that the work has not been used, neither completely nor in parts, for achieving any other academic degree.

Hohenheim / April, 2018

Manuel Vázquez Arellano