Christoph Schmied, BSc.

# Determination of Relative Orientation based on Distance Measurements

**MASTER'S THESIS**

to achieve the university degree of

Diplom-Ingenieur

Master's degree program: Geomatics Science

submitted to

**Graz University of Technology**

Supervisor

Ao.Univ.-Prof. Dipl.-Ing. Dr.techn. Manfred Wieser

Institute of Geodesy

Graz, March 2020

# Affidavit

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitly indicated all material which has been quoted either literally or by content from the sources used. The text document uploaded to TUGRAZonline is identical to the present master's thesis.

Date

Signature

# Acknowledgements

I would like to thank my supervisor Ao.Univ.-Prof. Dipl.-Ing. Dr.techn. Manfred Wieser, for his guidance through every aspect of this thesis, for his know-how and support. I would also like to thank him for giving me the possibility to write my thesis as a student project assistant in his team.

A huge thank you goes to Dipl.-Ing. Eva Reitbauer for all your advice, encouragement and support throughout this time. I really acknowledge our discussions on complicated and often times non-trivial topics. I also would like to express my gratitude for supporting me in generating ground truth data and assisting me during the measurement campaigns.

Special thanks go to Fachoberinspektor Bernd Mölg for helping me build the mechanical structure of the test setup. I also do owe him a debt of gratitude for his support during the measurement campaigns.

Moreover, I wish to show my appreciation to the company Pusch & Schinnerl GmbH. Without them, there would have been no impulse for writing this thesis in the first place.

I am deeply grateful for having wonderful and supportive classmates during my studies. Some of them became dear friends of mine over the years. Without them it would have not been the same.

Most of all I would like to thank my family for their love and overwhelming support throughout my academic career. Thank you for letting me do what I want and for backing me up in following my dreams.

Last but not least I want to thank the most important person in my life, my girlfriend Lisa. Thank you for always cheering me up when times are tough. Thank you for your love and always being there for me.

Thank you!

# Abstract

Composting is a growing industry worldwide and recently ideas emerged to automate the process of turning bio-waste into compost. If an automated compost turner should driven along these piles, also called windrows, it has to know it's position and attitude. The usual approaches of heading determination by using consecutive Global Navigation Satellite System (GNSS) measurements or by using an Inertial Measurement Unit (IMU) cannot be used on their own, because of the very slow speed of the machine and the heavy vibrations generated by it. Therefore this thesis deals with the determination of relative orientation in a challenging environment by utilizing novel methods based on distance measurements with regard to the triangle-shaped compost piles below the compost turning machine. As means of creating these measurements, ultrasonic and laser sensors, as well as a stereocamera have been evaluated. New approaches regarding the algorithms used to derive heading information from distance measurements have been developed. They range from simple approaches where the peak of a pile is estimated via line intersection to more sophisticated methods like surface profile correlation or Visual Odometry (VO). The results of these algorithms have been validated with two measurement campaigns where high-precision ground truth data was generated with the help of surveying. By utilizing ultrasonic or laser measurements, results have shown that the accuracy, precision and robustness has not been high enough to steer an automated compost turner alongside a windrow. Measurements from the stereo camera in combination with the surface profile correlation algorithm have proven to be accurate enough over small periods of time. The RMSE of relative heading for a 15 minute drive was as low as 4 degrees. With the help of more sophisticated methods, like VO a very low RMSE of 0.6 degrees over the course of a 20 minute drive alongside a windrow has been achieved. The results of this thesis show, that image-based navigation techniques can be used in agricultural applications with harsh environmental conditions.

# Kurzfassung

Das starke Wachstum der Kompostierungsindustrie führte in letzter Zeit zu Überlegungen hinsichtlich der Automatisierung des kompletten Kompostierungsprozesses. Die wichtigste Information für einen automatisierten Kompostwender ist seine Position und Ausrichtung. Normalerweise können all diese Information durch Technologien wie GNSS oder inertiale Messeinheiten geliefert werden. Durch die sehr langsame Fortbewegungsgeschwindigkeit und die starken Vibrationen ist dies in diesem Fall nicht möglich. Diese Arbeit behandelt alternative Technologien zur relativen Richtungsbestimmung von Kompostwendemaschinen im Bezug auf industrielle Dreiecksmieten. Aus Distanzmessungen zu diesen Mieten, welche mithilfe von Ultraschall-, Laser- oder bildgebenden Sensoren getätigt werden, werden im Anschluss Richtungsinformationen abgeleitet. Neuartige Algorithmen wurden hierbei untersucht. Zu Beginn wurden einfache Methoden analysiert, wie die Herleitung von Richtungsinformation durch Schneiden zweier Geraden, welche entlang der Oberfläche der Dreiecksmiete verlaufen. Des Weiteren wurden komplexere Methoden analysiert, bei welchen durch die Korrelation von Oberflächenstrukturen Richtungsinformation abgleitet werden kann. Auch hochkomplexe Methoden zur relativen Richtungsbestimmung, wie Odometrie durch bildgestützte Sensorik, wurden betrachtet. Distanzmessungen wurden während zweier Messkampagnen aufgenommen und die Resultate diese Messungen in Kombination mit den Algorithmen wurden mithilfe von Referenzmessungen validiert. Die Verwendung von Ultraschall- oder Lasermessungen erzielte keine brauchbaren Ergebnisse mit ausreichender Genauigkeit oder Präzision um einen autonomen Kompostwender steuern zu können. Die Verwendung von Distanzmessungen einer Stereokamera in Kombination mit einer Oberflächenkorrelationsmethode lieferte gute Ergebnisse für kurze Zeitintervalle kleiner 15 Minuten. RMS-Fehlerwerte um 4 Grad über diese Zeitspanne wurden erzielt. Die Methode der bildgestützten Odometrie lieferte ausgezeichnete Ergebnisse. Diese Methode zeichnet sich durch Effektivwert-Fehler von 0.6 Grad über die komplette Dauer einer Fahrt von 20 Minuten aus. Die Resultate dieser Arbeit haben gezeigt, dass sich bildgestützte Methoden zur Navigation sehr gut für die Anwendung im Agrikulturbereich eignen.

# Acronyms

**BLE** Bluetooth Low Energy.
**BRDF** Bidirectional Reflectance Distribution Function.

.

**DME** Distance Measurement Equipment.

**EMI** Electromagnetic Interference.

**FMCW** Frequency Modulated Continuous Wave.
**FOV** Field of View.

**GMT** Greenwich Mean Time.
**GNSS** Global Navigation Satellite System.
**GPS** Global Positioning System.
**GPST** Global Positioning System Time.

**I2C** Inter-Integrated Circuit.
**IMU** Inertial Measurement Unit.

**MAD** Median Absolute Deviation.
**MCU** Microcontroller Unit.
**MMAD** Moving Median Absolute Deviation.

**PCB** Printed Circuit Board.

**RANSAC** Random Sample Consensus.

**SDK** Software Development Kit.
**SLAM** Simultaneous Localization and Mapping.

**TAI** Temps Atomique International (International Atomic Time).
**TDM** Time-Division Multiplexing.
**TOF** Time of Flight.

**UAV** Unmanned Aerial Vehicle.

**UT** Universal Time.

**UTC** Universal Time Coordinated.

**VO** Visual Odometry.

# Contents

# List of Figures

# List of Tables

# Introduction

The introduction tries to give a general overview of the Master thesis *Determination of Relative Orientation based on Distance Measurements*. In the following sections, the motivation behind this thesis will be discussed and research goals will be defined. To wrap up this chapter, the execution of this thesis will be summarized in the last section of the introduction.

## Motivation

Due to the everlasting growth of the economy and the arising waste problems, it is even more important today to recycle waste products and get them to good use. Biodegradable waste in particular has a huge potential, because it can be composted and given back to nature as fertilizer to complement soil for future generations. Unfortunately, the task of creating compost, is anything but straightforward. The biodegradable material is piled up in huge triangle-shaped lots, called windrows. [1] To get a working composting process, it is necessary to mix the media periodically with the appropriate amount of water and oxygen (air). At composting facilities, this is usually done with huge human-controlled, diesel-engine powered machines. The labour conditions at aforesaid waste treatment facilities are very harsh. Due to this, the idea emerged to automate the task of inter-mixing the biodegradable material. Nowadays, automated driving is nothing new and several standards have been developed to handle all sorts of situations and problems. These compost turning machines, on the contrary, brought up some difficulties regarding the process of autonomous driving, which have not been addressed or solved yet. One of the difficulties is determining the relative orientation of the machine with respect to the windrow. At first, this problems seems to be easy to solve with modern technologies like high-end Global Navigation Satellite System (GNSS) receivers. Similar problems have also been solved when combining GNSS measurements with accelerometer and magnetometer data. However, when inspected further, many difficulties arise. First and foremost, due to

the extremely slow speed of the machine (0.3 km/h), it is challenging to acquire viable heading information from successive GNSS measurements. Because of the vibrations generated by the diesel engine, the signal-to-noise ratio of the IMU's measurements is very low. The fact that automatizing the composting process could lead to larger bio fertilizer production and thus motivate people to grow organically, is the strongest motivation for this thesis.

## Research goals

The ambition of this thesis is to acquire the relative orientation of the compost turner by utilizing Distance Measurement Equipment (DME) aimed at a triangle-shaped windrow. Three types of DME will be evaluated regarding their accuracy, precision and reliability. Besides, various algorithms for heading determination, based on the aforesaid method, will be examined. The analysis of a windrow's surface profile will also play a key role in some of the algorithms considered.

The focus is to develop a proof-of-concept, that measurements from DME can be taken and further processed into a solution where relative orientation of the vehicle is estimated precisely.

## Execution

The thesis is divided into three main parts. Part I addresses the concept of relative orientation, the measurement principles of the sensors and the theory behind various algorithms. Part II will deal with the practical test setup, investigations regarding accuracy and precision of the sensors and the two measurement campaigns carried out during the tests. In Part III the results of those campaigns will be presented. This part will also try to compare the practical implementation of the algorithms analysed in Part I. At last, we will discuss what was achieved during the work in this thesis and give an outlook on further development in this field.

# Status quo

This chapter will give a brief overview of past and current development in the field of this thesis. Especially, the present state of autonomous track-driven vehicles will be outlined. In addition to this, a general view on the composting industry in Europe, notably Austria, is made.

## Autonomous track-driven vehicles

Track-driven vehicles in everyday life have been most established in the sector of heavy machinery. Huge excavators or snow groomers are almost always track-driven. This is due to the perfect weight distribution of tracks, when compared to wheels. The weight transfer of wheel-driven machines can mostly be described as ”point-loading”, whereas with track-driven vehicles the mass of the machine is very well distributed. The properties of tracks really stand out when soft ground conditions are considered. Stability is another crucial factor, notably on slippery ground. Because of these factors, most of the development regarding autonomous track-driven vehicles in history has been done in the sector of space exploration in the form of rovers [2]. The second most investigated topic in literature is the integration of autonomous machinery in construction sites [3]. Another interesting topic is the development of robotic snow-groomers [4], which could optimize the grooming operation in skiing regions. One important factor for track-driven vehicles is the detection of slip events of the tracks. Slip can be described as the difference between the track velocity and the forward velocity of the vehicle. Therefore, odometry data has to be integrated with other sensor data, like GNSS measurements, which are not influenced by slip. Investigations on autonomous track-driven compost turners cannot be found in literature as of yet.

# Composting industry

The European Compost Network [5] states, that roughly 118 to 138 million tonnes of biodegradable waste get generated in the European Union annually. Only 25% of this huge mass get recycled and turned into valuable high-quality compost. The major part of the organic waste is still put into landfills across Europe. Since more and more farmers begin to switch from commercially available chemical fertilizer to organic alternatives, the demand for compost is on the rise. The benefit of this material is, that recurring additions to soil create a better structure, reduce erosion and improve fertility. All these factors boost crop yield. The current way-of-work of composting facilities (Figure 1) is, the composting of large quantities of materials piled up in windrows. These windrows get turned and aerated by a human-controlled compost turner. This process can be very tedious and dull for the operator. Therefore, it is difficult for compost plant operators to find personnel. A new and modern approach (i.e. autonomous compost turners) could lead to an increase in availability for bio-fertilizer and therefore to a more sustainable community overall.

Figure 1: Windrows of a composting facility in Alberndorf in der Riedmark near Linz, Austria

# Part I

# Theoretical foundations

# 1 Sensor technology

This chapter aims at explaining different types of technologies regarding distance measurement. Each distance measuring technology has its advantages and drawbacks. Therefore, the best sensor technology has to be specifically selected for each application. The properties, as well as the theoretical accuracy and precision of each sensor type, will be discussed. Moreover, the robustness in harsh environmental conditions will be investigated.

## 1.1 Research on potential technologies

The main goal of this thesis is the determination of relative orientation of a compost turner with respect to a windrow. The idea is to utilize distance measurements to determine the surface profile of the windrow below a compost turning machine. Afterwards it will be validated, if the differences of concurrently generated surface profiles is significant enough to derive heading information.

First and foremost, the most reliable methods of distance measurement, with respect to the material compost, have to be evaluated. Due to the uneven shape and granularity of the scattered media, it is sufficient to have a distance accuracy in the range of centimeters. The detection of the surface profile may require a large quantity of sensor components, and thus price can also be an economically important factor for the realization of this concept.

Conditions for Distance Measurement Equipment (DME) at a composting plant can be very harsh and also challenging. During the turning process, the material radiates away huge amounts of heat and steam. This can be an immense problem for certain types of distance measuring technologies. Moreover, the rotational speed of the spiked drum leads to high dust formations and also larger chunks of material getting dispersed through the air. In addition to that, windrows can have various textures and reflective properties, complicating reliable distance measurements even more.

Summarized, this leads to the following requirements in the selection of an appropriate technology:

- Reliable measurements regarding the material compost;
- Accuracy in the range of centimeters;
- Undisturbed by environmental conditions (temperature, humidity);
- Mid-to-low price range per sensor;
- Dust particles should have little to no influence on the measurement.

In-depth analysis has shown, that most of the established technologies on the market had at least one drawback regarding the aforesaid requirements. Four measurement technologies have been selected and evaluated further. An overview of the specification and possible draw backs is shown in Table 1.1

| Type | Accuracy | Range | Price | Principle | Disturbance |
|---|---|---|---|---|---|
| Ultrasound | + | + | ++ | TOF | Dependent on temperature and humidity |
| Radar | ++ | +++ | - | TOF, FMCW | EMI; possible dead zones |
| Laser | +++ | ++ | - | TOF | Problems in dusty environments |
| Stereo camera | ++ | + | o | Stereo Vision | Pictures of compost can be homogeneous |

Table 1.1: Specification of evaluated sensor technologies; [+] requirement fulfilled, [-] requirement not met, [o] neutral
Time of Flight (TOF)
Frequency Modulated Continuous Wave (FMCW)
Electromagnetic Interference (EMI)

Ultrasonic sensors in the mid-to-low price range category fulfill all the stated requirements regarding accuracy, range and price. The downside of this technology is that the measurements heavily depend on temperature and humidity (subsection 1.2.1). However, this dependency can be mitigated by correcting the measurements with in-field collected environmental data. Another ambiguity is the reflective index of sound waves regarding the media compost. This will be elaborated further in section 4.1.

Radar, at first glance, seems to be the best-fitted technology for this thesis. It is sovereign regarding dust particles and dirt in the air. For this reason, it is also used in monitoring the level of media in circular silos. [6] Due to the high resolution, these sensors can deliver very accurate distance measurements. The only thing to consider is the dead-zone of the radar sensor used. A dead-zone describes the minimal distance

a radar sensor can reliably measure. Radar sensors which fulfill all the requirements lie in the upper price range and thus it will not be feasible to use a large quantity of those.

By using laser sensors for surface determination, one can get a very detailed representation of it. The detail, modern laser scanners can achieve, is outstanding, but would not fit the intended budget of this thesis. Laser sensors in the right price range do have some major drawbacks when it comes to reliability and failure safety in challenging and dusty environments.

In the last years, stereo vision has seen a huge boost in market share regarding navigation of robots, especially UAVs. By using the principle of stereo vision for navigation, in combination with sophisticated VO algorithms, it is possible to acquire very accurate and reliable pose estimations. When combining the results with measurements from GNSS receivers and IMUs in a Kalman-Filter [7], one gets a very fail safe system.

### 1.1.1 Selection of suitable sensors

During the research for suitable sensory technology, three of the four systems mentioned in section 1.1 have been selected for evaluation in this thesis. Radar sensors have been dropped, because of the low availability in the low-to-mid-range price segment. In Table 1.2 the comparison of the selected hardware components is shown.

| Type | Name | Max. range | FOV | Resolution | Update rate | Price* |
|---|---|---|---|---|---|---|
| Ultrasound | HCSR-04 | 4 m | 30° | 5 mm | 50 Hz | 2-5 $ |
| Laser | VL53L1X | 4 m | 27° | 1 mm | 30 Hz | 15 $ |
| Stereo camera | ZED | 20 m | 110° | 1 mm | 15 - 100 Hz | 450 $ |

Table 1.2: Selected DME hardware for later evaluation; *Price per sensor

## 1.2 Ultrasonic distance measurement

This section will give a brief overview of the fundamentals of ultrasonic ranging.

The base principle of ultrasonic ranging is two-way Time of Flight (TOF). TOF is the time taken by a wave (electromagnetic or acoustic) to travel through a medium. An ultrasonic distance sensor uses this phenomenon by sending out a short acoustic pulse

and waiting for the response. The time it takes for the wave to travel to the object and back is double the distance from the sensor to the object, therefore two-way TOF. Acoustic waves have different characteristics when propagating in various materials and media. These characteristics have to be considered when utilizing aforesaid waves for distance measurement and will be approached in subsection 1.2.1.

### 1.2.1 Speed of sound

An acoustic wave is a longitudinal wave, which propagates with the speed of sound. At 20°C, the speed of sound through air is roughly 343 meters per second. The rate of change in velocity is strongly dependent on temperature and humidity. The medium in which sound waves travel is also an important factor to consider. However, in most cases the medium will be air. The basic formula for the speed of sound is

$$c = \sqrt{\gamma \cdot \frac{p}{\rho}} \tag{1.1}$$

where

- $\gamma$ is the adiabatic index (ratio of the specific heat of gas at constant pressure to that at constant volume);
- $p$ is the pressure;
- $\rho$ is the density.

When using the ideal gas law, Equation 1.1 becomes

$$c_{ideal} = \sqrt{\frac{\gamma \cdot R \cdot T}{M}} = \sqrt{\frac{\gamma \cdot k \cdot T}{m}} \tag{1.2}$$

where

- $R$ is the molar gas constant;
- $M$ is the molecular weight of the gas;
- $k$ is the Boltzmann constant [8];
- $T$ is the absolute temperature;
- $m$ is the mass of a single molecule.

Substituting all the before mentioned constants with numeric values, the speed of sound in air is

$$c_{air} = 331.3 \cdot \sqrt{1 + \frac{\vartheta}{273.15}} m/s \tag{1.3}$$

where $\vartheta$ is the temperature in degrees Celsius.

Ultimately, Taylor expansion of the remaining square root in $\vartheta$, yields the practical formula for the speed of sound in dry air:

$$c_{air} = (331.3 + 0.606 \cdot \vartheta)\ m/s \tag{1.4}$$

As stated previously, Equation 1.4 is only valid in dry air (0% relative humidity). When humidity is considered, Equation 1.2 has to be analyzed further. The adiabatic index $\gamma$, as well as the molecular weight of a gas $M$, will change depending on the level of humidity. The adiabatic index $\gamma$ is calculated by

$$\gamma = \frac{c_p}{c_v} \tag{1.5}$$

where

- $c_p$ is the heat capacity at constant pressure;
- $c_v$ is the heat capacity at constant velocity.

The heat capacity of humid air can be computed by:

$$c_p = 1.005 + 1.82 \cdot H$$
$$c_v = c_p - \frac{R}{M}$$

where 1.005 kJ/kg °C is the heat capacity of dry air, 1.82 kJ/kg °C the heat capacity of water vapor, and $H$ is the absolute humidity [9] in kg water vapor per kg dry air in the mixture. Due to the fact that the molecular weight of humid air is less than the weight of dry air, humid air is less dense. Thus the speed of sound in humid air is greater than the speed of sound in dry air.

The effect of temperature and humidity on the speed of sound is visualized in Figure 1.1. The growing importance of humidity with rising temperatures is clearly evident. E.g. the error for a 20 ms TOF pulse between 0% RH and 100% RH at 80 °C would be 0.5 m.

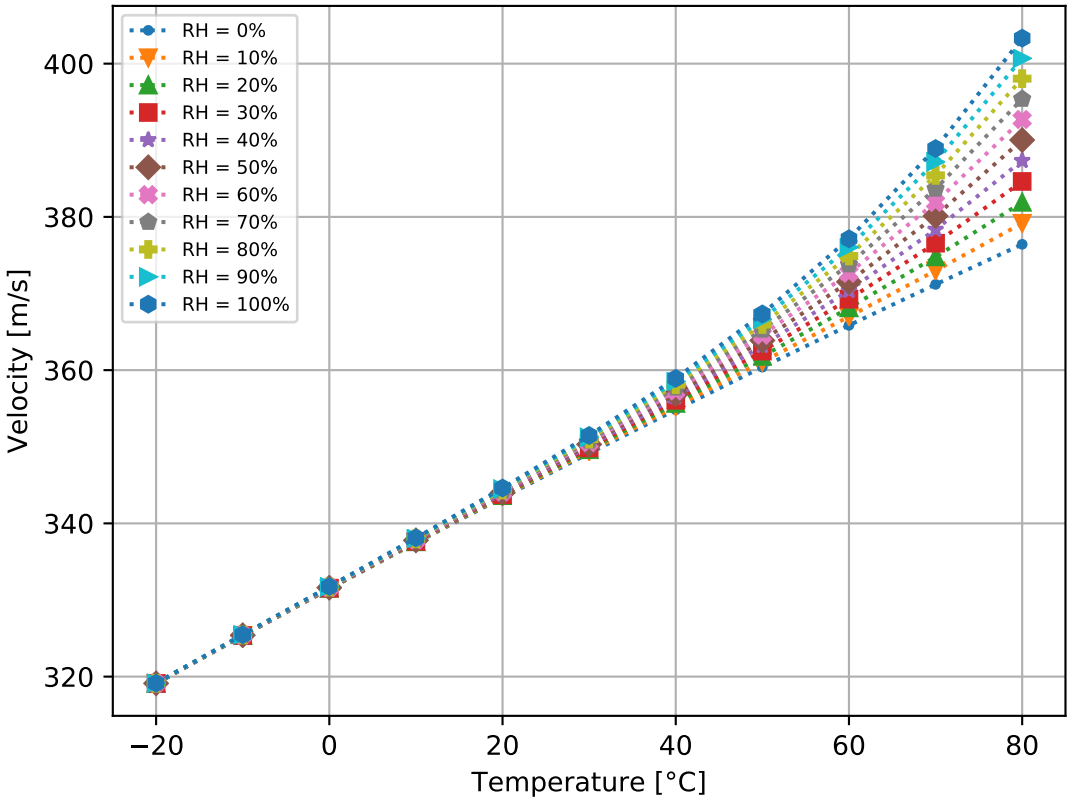



Figure 1.1: 100 °C temperature sweep over various humidity levels

To sum up this section, one can assume that the speed of sound in a medium is inversely proportional to the square root of its density and is heavily dependent on temperature. The density of air is correlated with relative humidity, but the effect of changing density on the speed of sound is much more insignificant, although it gets progressively more distinct with increasing temperature, as seen in Figure 1.1.

### 1.2.2 Signal characteristics

The usual range of frequency human beings can hear, is between 20 Hz and 20000 Hz. Anything over 20 kHz is defined as ultrasound. A typical ultrasonic DME uses a frequency starting at 40 kHz up to 100 kHz. The signal usually consists of a sequence of pulses and is illustrated in Figure 1.2.

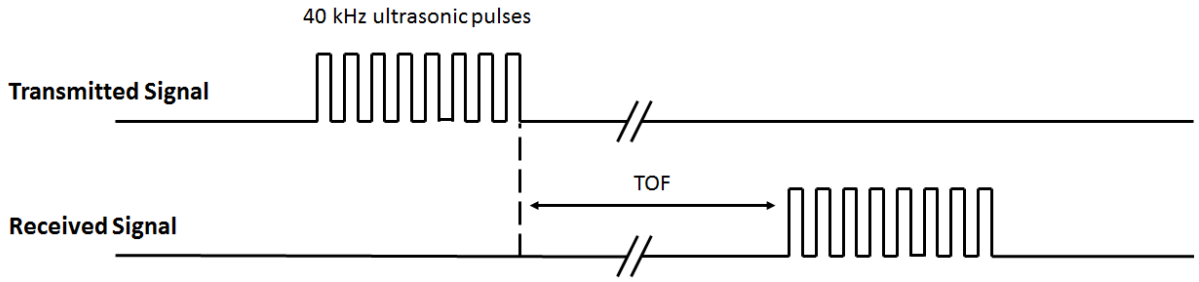



Figure 1.2: Typical ultrasonic ranging signal

More sophisticated ultrasonic DME use a carrier wave, in combination with a modulated signal, to estimate the length of the signal in the time domain. These are also able to detect multiple echos in a single measurement cycle. Low-cost alternatives only detect the first echo and start a new measurement cycle after that. This behaviour is visualized in Figure 1.3.

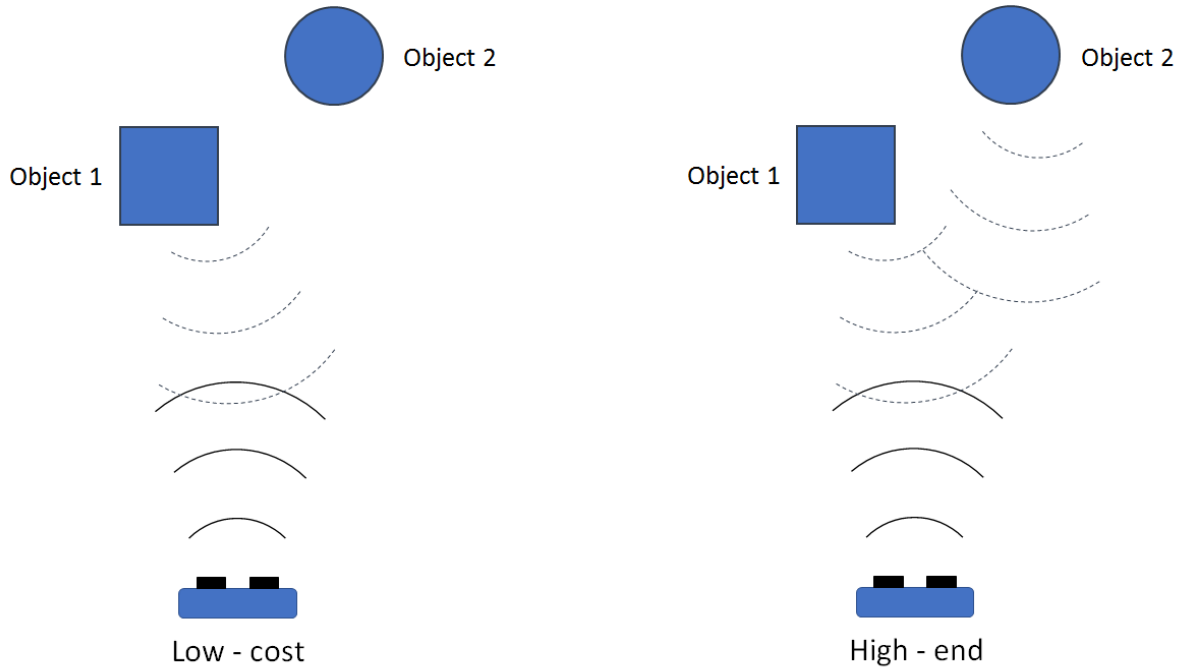



Figure 1.3: Comparison of low-cost and high-end ultrasonic DME

Summarized, this means that the quality of the measurement, regarding signal characteristics, depends on the frequency as well as the modulation technique used. By only using the first echo, it is possible to have a higher sample-rate, with the disadvantage of only recognizing the closest objects. Multiple echo DME, on the contrary, are more suitable for surface recognition.

## 1.3 Laser distance measurement

The working principle of most laser distance measurement sensors is very similar to that mentioned in section 1.2, as it is also TOF. The main difference between the two is the characteristics of the emitted signal and therefore the wavelength of the signal. Ultrasonic distance measurement depends on the emission of short ultrasonic bursts, whereas in laser distance measurement intense focused beams of light of a single frequency get emitted. The greater frequency of the laser beam, leads to a better resolution in the received echo of the signal, and thus to better accuracy. Furthermore the update-rates of laser sensors are higher, because the speed of light is magnitudes above the speed of sound.

Difficulties can arise, if too much background light interferes with the laser pulse. In such conditions, the receiving optics of the laser sensor can falsely detect background light as the reflected laser pulse, resulting in spurious readings. These noise conditions can be mitigated by using narrow-bandwidth or split-beam frequencies [10] to block out interference from background light. Another drawback of this technology is the vulnerability in dusty environments, as stated by [11].

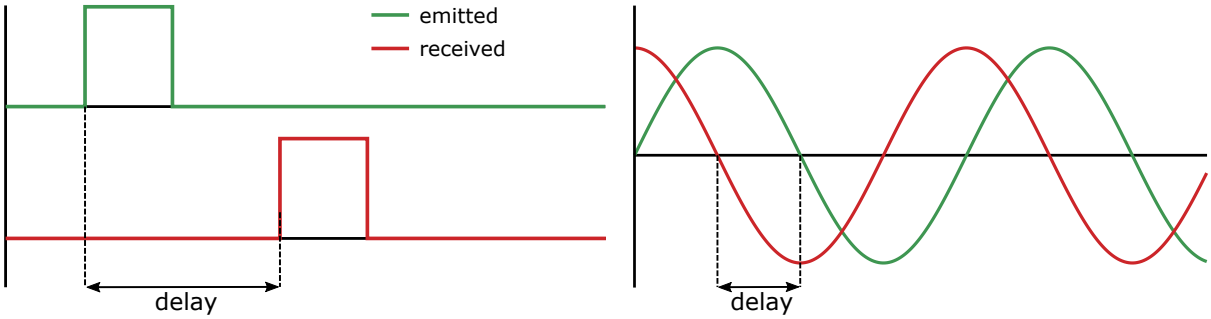



Figure 1.4: Left plot: *Direct method*, based on the time difference of transmitted and received pulse; Right plot: *Indirect method*, based on the measurement of the phase difference of two signals

The method described before is also often mentioned as the *direct* TOF method, as it measures the time difference between the emitted and a received signal. Another method of distance measurement using lasers is called the *indirect* method. In this approach, a continuous modulated sinusoidal light wave is emitted and the phase difference between outgoing and incoming signals is measured. Both of these methods are visualized in Figure 1.4.

## 1.4 Distance measurement by stereo vision

This section will describe how to extract distance information by utilizing stereo vision principle.

The geometry of a binocular stereo vision system is shown in Figure 1.5. In its most simple form, this system consists of two identical cameras only separated by a fixed distance in x direction, the baseline $B$.

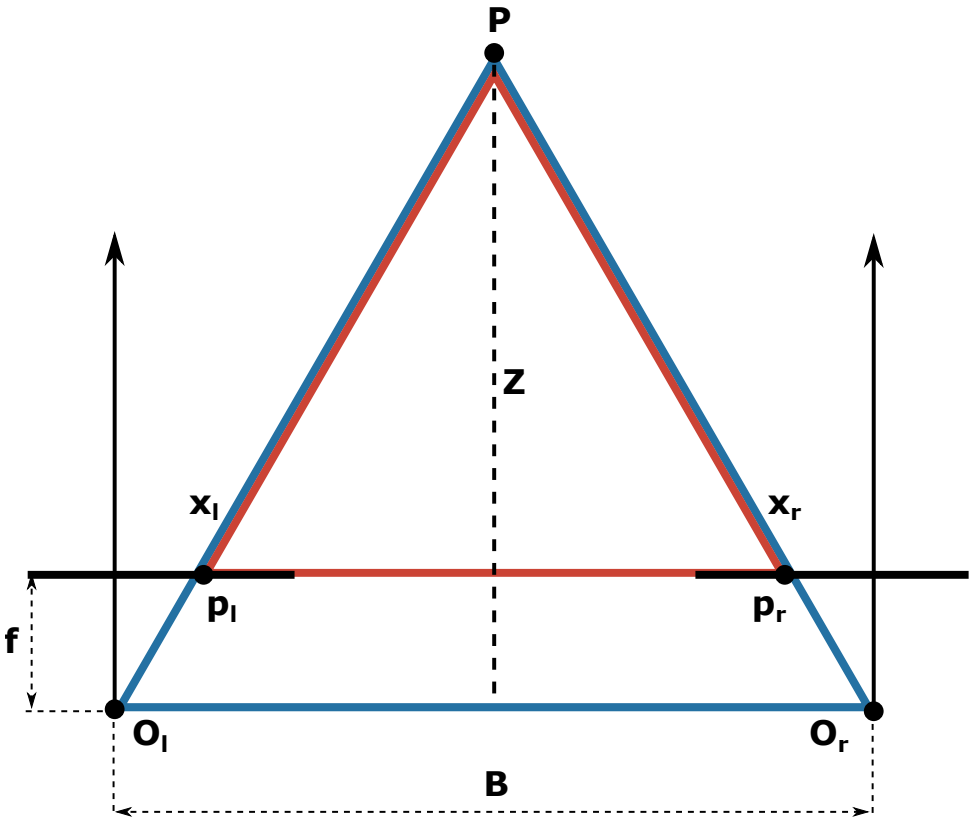



Figure 1.5: Model of parallel stereo camera

In this form, the image planes (bold black lines) are co-planar. A point in real-life $P$ is viewed by the two cameras at different positions in the image plane ($p_l$ and $p_r$). The displacement of the two points is called *disparity*. By looking at the triangles $PO_lO_r$ (blue) and $Px_lx_r$ (red), one can derive that

$$\frac{B}{Z} = \frac{B + x_l - x_r}{Z - f} \tag{1.6}$$

and therefore compute the depth $Z$ by using

$$Z = f\frac{B}{x_r - x_l} = f\frac{B}{d} \tag{1.7}$$

where $f$ is the known focal length and $x_r - x_l$ is the disparity $d$. This means, that the depth for a specific pixel in the image can be recovered by knowing the disparity at this pixel. To know the corresponding coordinates $x_l$ and $x_r$, one has to identify conjugate pairs in the images. The problem in detecting those pairs is also known as the *correspondence problem* and is a non-trivial problem to solve. To locate the matching points, stereo vision utilizes edge and region feature detectors. Due to the epipolar constraint [12], the search space for these features is limited to the pixels of a horizontal line in the image. However, measurement errors and uncertainties in camera position and orientation can lead to conjugate pairs which do not fulfill this constraint exactly. Is this the case, a small neighborhood of pixels can be considered in feature matching algorithms.

To summarize, the acquisition of depth values by stereo matching two images consists of the following steps:

1. Calibration of both cameras
2. Rectification of the images to fulfill the epipolar constraint
3. Feature matching
4. Disparity calculation from matched features
5. Derive depth values from disparity map

# 2 Algorithms for heading determination

In this chapter, various algorithms for obtaining orientation information from distance measurements to triangle-shaped objects are introduced. The assortment of algorithms range from very simple concepts, like finding shift in position, to much more sophisticated ideas. The basic principle behind all the concepts is the determination of features, which describe the current position of the compost turning machine with respect to the windrow below it. From the difference in position of the features in consecutive epochs, valid heading information will be derived.

## 2.1 Simple approach based on line intersection

The most simple approach of getting features, which describe a shift in position, is based on the intersection of two lines estimated from points on the slopes of a windrow. The point of intersection will be the feature to track in consecutive epochs. The coordinates of the sensors in the xz-plane of the compost turning machine have to be known. By using these coordinates, in combination with the distance measurement and the alignment of the sensor, sampling points with $x$ and $z$ coordinates for the lines can be computed. A basic graphical overview of this concept is given in Figure 2.1.

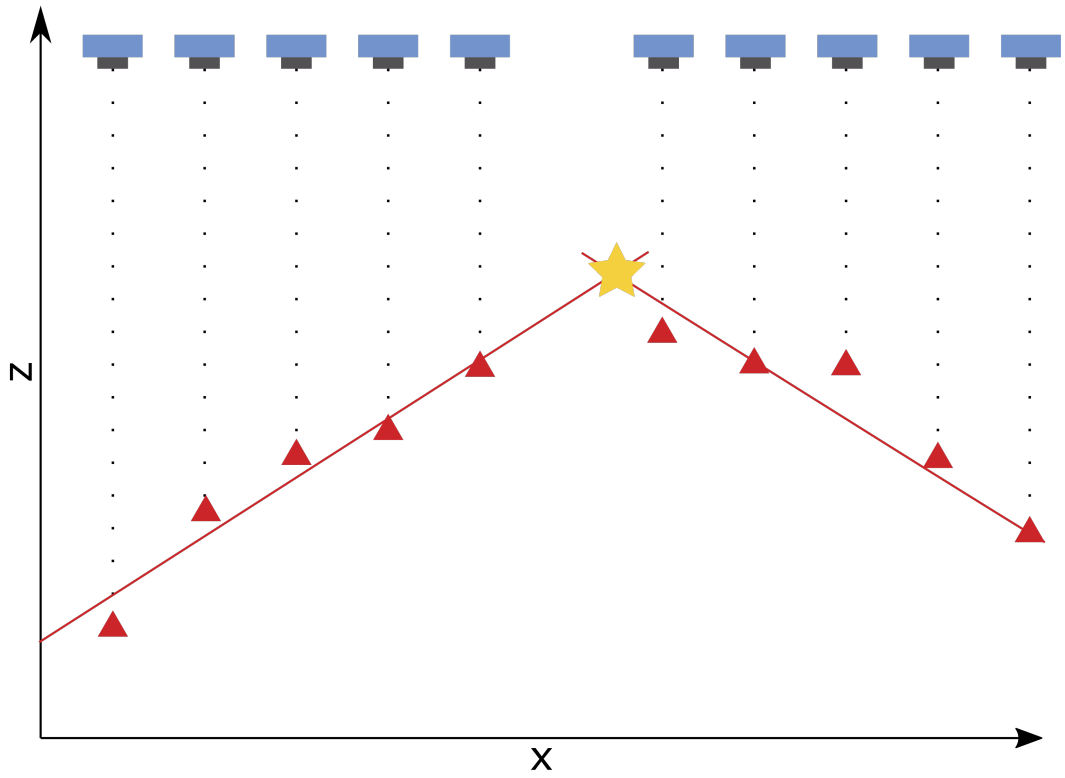



Figure 2.1: XZ-plane view of sensor configuration; ultrasonic sensors in top area (blue); sampling points (red); intersection point or peak of windrow (yellow)

At first, the measurements get separated into two groups. The first group contains the measurements left of the peak of the window, the other group the measurements right of the peak. The distance measurements $r$ and the arrangement angle[1] of the sensor $\alpha$ can then be transformed into sampling points with $x$ and $z$ coordinates by

$$x = x_{sens} + r\cos\alpha \quad and \quad z = z_{sens} + r\sin\alpha, \tag{2.1}$$

where $x_{sens}$ and $z_{sens}$ are the coordinates of the sensors. The next step is to estimate the two lines from which the intersection point shall later be derived. This can be done by a simple least-squares adjustment or other more robust methods like Random Sample Consensus (RANSAC) [13]. For further calculation, only the x-coordinate and y-coordinate of the current detected peak will be used. The y-coordinate of the peak is estimated from the speed of the vehicle in combination with known time intervals between measurements. By using

$$y_t = y_{t-1} + v \cdot \Delta t, \tag{2.2}$$

[1] tilt-angle of the sensor with respect to the z-axis of the bodyframe

where $v$ is the speed of the vehicle and $\Delta t$ the time between measurement epochs, the relative y-coordinate in relation to a starting point $y_0$ can be computed. The relative heading $\gamma$ (denoted as $\alpha$ in Figure 8.3) can then be calculated by using

$$\gamma = \arctan\left(\frac{y_t - y_{t-1}}{x_t - x_{t-1}}\right). \tag{2.3}$$

When the speed of the vehicle is very slow and thus the difference in y-coordinate small, longer time intervals $\Delta t$ shall be used to get reliable heading information.

## 2.2 Surface profile extraction

This variation of feature extraction from range measurements tackles the determination of a complete surface profile. The methods used in this approach can handle a wide range of sensor configurations and synthetic arrays, realised by the movement of several sensors. Therefore, this algorithm is highly flexible regarding the arrangement of sensors. A very detailed explanation of this technology can be found in [14].

The basic concept is, that the coordinates of the sample points, obtained by for example section 2.1, get extended to an array of arcs containing multiple sample points per measurement. The opening angle of this arc depends on the beam-width of the utilized sensor. Thus, a variety of range measurements from sensors with known beam-widths can be used. For ultrasonic receivers, there are two main categories of sensors. The first category, seen in Figure 2.2a, consists of separate transmitting and receiving elements and therefore the resulting uncertainty of the measurement lies on an elliptical arc. In Figure 2.2b, one transceiver serves as transmitter and receiver and the uncertainty has the form of a circular arc. In this visualization, $\theta$ implies the beam-angle of the sensor and $r$ the measured distance. By utilising this technology, we do not have the constraint of only gaining one sample point along the line of sight. Instead we get an arc, representing the uncertainty of the object's location. Close to actual reflection points of the surface we want to profile, several arcs will coincide and many small arc segments get superimposed. These regions can then be analyzed and processed further using one of the techniques shown in subsection 2.2.1 and subsection 2.2.2. Both of these techniques create an area of xz-coordinates and treat this area like an ordinary picture in which the arc points are represented as pixels.

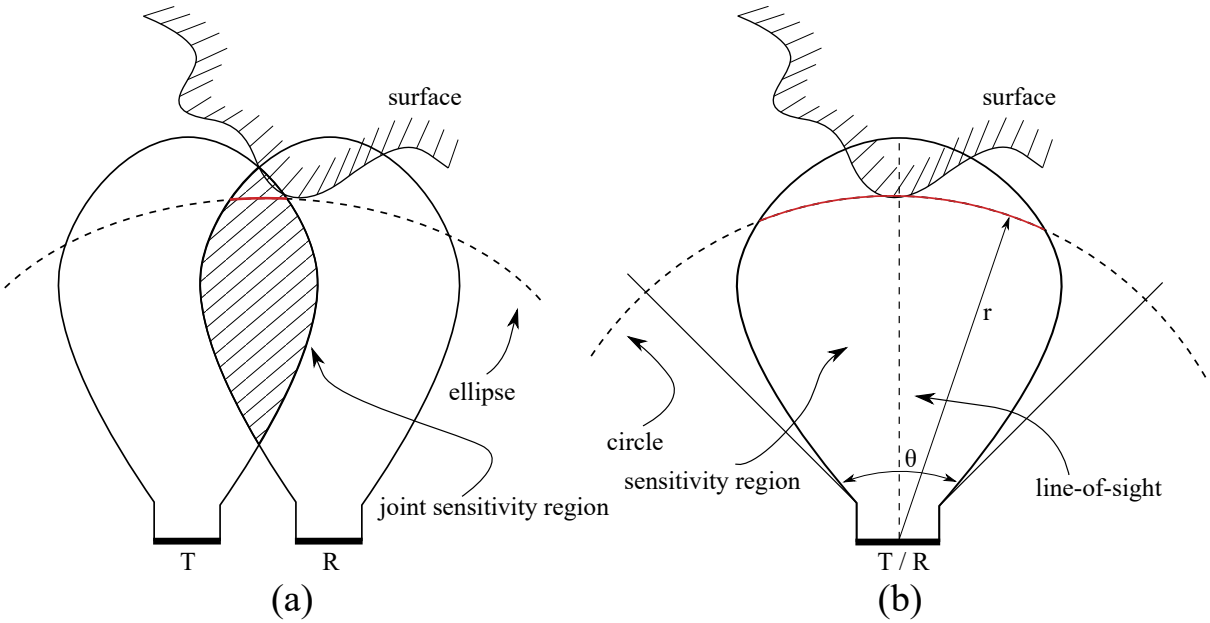



Figure 2.2: (a) An ultrasonic transducer-pair, where the arc response has an elliptical form and a small joint sensitivity region. (b) Transducer serves as both, transmitter and receiver; the arc response is circular.

To summarize, the utilization of multiple range measurements in combination with successive arc map processing has several major improvements over the simple approach described in section 2.1.

### 2.2.1 Spatial voting

One way of processing the arc map is to apply a spatial voting scheme followed by thresholding. The basic idea is to track the number of pixels where arcs are superimposed. These overlays can then be summed up and represent the pixel value of an image, where x and z coordinates imply the pixel's location. The most important parameter of the spatial voting scheme is the threshold value. This value has to be chosen adequately, so that only pixels which lie above it are considered valid. The valid pixels are then subsequently used as sampling points for the estimation of the true surface profile. In order to reliably estimate the surface of an object with this algorithm, a large number of measurements has to be present. The best results can be achieved by utilizing many sensors with various alignments. In Figure 2.3 an illustrative example of the spatial voting scheme can be seen. The more arcs superimposing in one pixel, the greater the pixel's value (represented as hue of red) and the more likely it will be, that this pixel is above a given threshold.

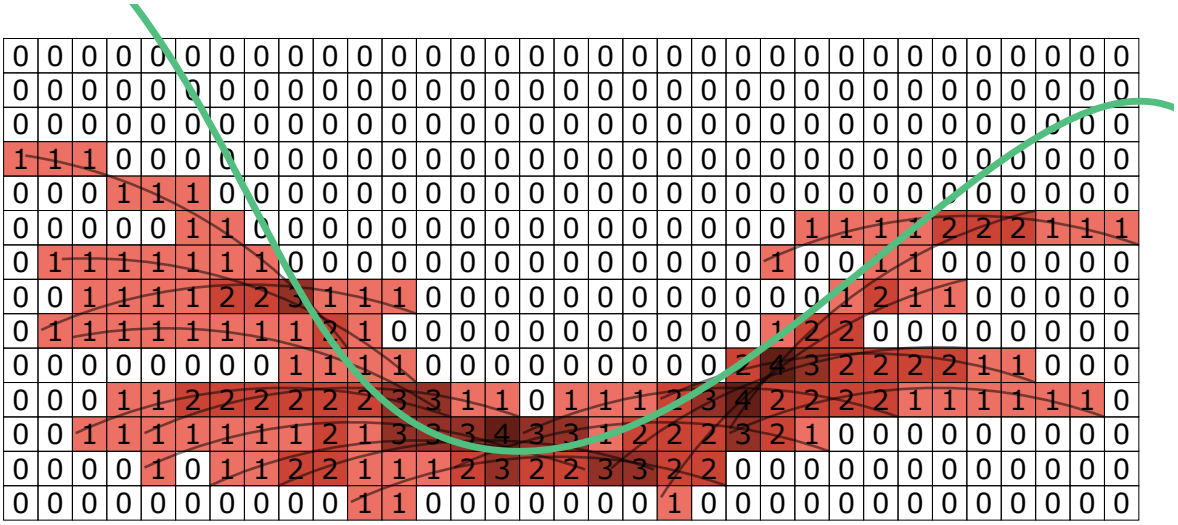


Figure 2.3: Spatial voting scheme pixel representation of example scene with measurement uncertainty arcs (black) and the true surface representation (green); Pixels contain the sum of the number of arcs which coincide with them; pixelated surface profile approaches real profile with increasing number of measurements

Due to this example, it is easy to grasp, that a large amount of measurements yields a better distinction between the actual surface profile and spurious readings of the sensors.

### 2.2.2 Morphological processing

As an alternative to spatial voting, morphological processing of the resulting arc map can be used. Morphology is a set of image processing tools, which is broadly used for operations like edge detection, segmentation and texture analysis. [15] In such an operation, the value of each pixel in the output is based on the input pixel and it's neighborhood. The most basic operations are called erosion and dilation and all further processing tools are a combination of those two.

#### Erosion

The mathematical definition of erosion, first of the two fundamental operations, is

$$A \ominus B = \{z \in E | B_z \subseteq A\} \tag{2.4}$$

where $A$ is a binary image in the Euclidean space $E$ and $B_z$ is the translation of the structuring element $B$ by the vector $z$. The erosion $A$ by $B$ is also often written as

$$A \ominus B = \bigcap_{b \in B} A_{-b}. \tag{2.5}$$

This means that only when the structuring element $B$ is completely contained inside $A$, the pixel values stay the same. Otherwise they are eroded. A visual representation of this concept can be seen in Figure 2.4.

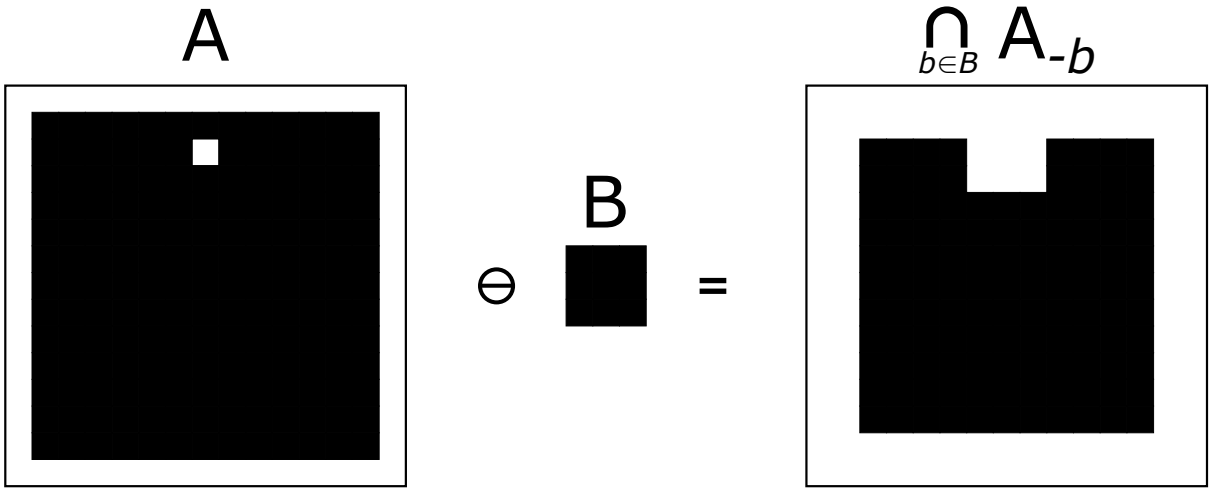



Figure 2.4: Pixels in $A$ get eroded by kernel $B$; Image is *thinned out*

**Dilation**

Dilation is the mathematical opposite of erosion and defined as

$$A \oplus B = \{z \in E | (B^s)_z \cap A \neq \emptyset\} \tag{2.6}$$

where $B^s$ stands for the symmetric of $B$. $A$ dilated with $B$ is also written as

$$A \oplus B = \bigcup_{b \in B} A_b. \tag{2.7}$$

This means that, for each pixel in $A$ with a value of 1, $B$ is superimposed with it's center on the pixel in $A$. In the output of the operation every superimposed pixel of $B$ is considered. This operation is visualised in Figure 2.5 and yields to a widening of small structures in binary images.

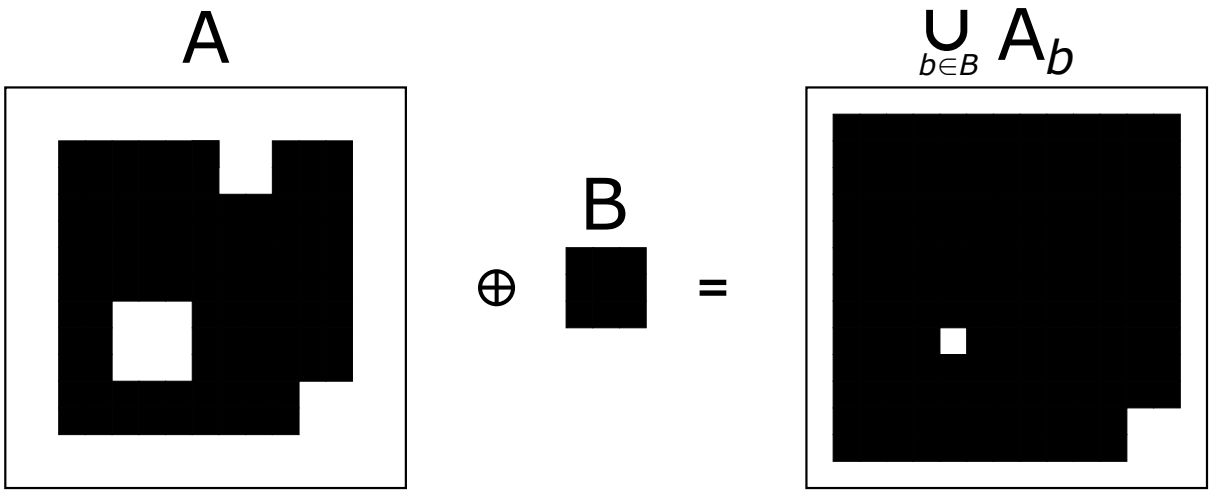



Figure 2.5: Pixels in $A$ get dilated by kernel $B$, yielding a *fattened* image

For the purposes of this thesis another important morphological operation, called *thinning*, is considered. When thinning is applied to a binary image, regions of equal-valued pixels get reduced to lines. The purpose of thinning is to scale back the image components to their essential information for further analysis and recognition. The mathematical definition of thinning is

$$A \otimes B = A - (A \circledast B), \tag{2.8}$$

where $(A \circledast B)$ is the Hit-and-Miss Transform [16]. Analogous to erosion and dilation, the principle of thinning is also based on the overlay of a structuring element's center of every pixel in a binary image. As a matter of fact, thinning can be seen as a variant of erosion, in which not all pixels of the structuring element have to match with the image, for the operator to be valid. The only difference in thinning is, that one has to specify a *hit* and a *miss* structuring element. The pixel of the image only gets modified, if all pixels of the *hit* and none of the *miss* structuring elements match. This operation is visualized in Figure 2.6.

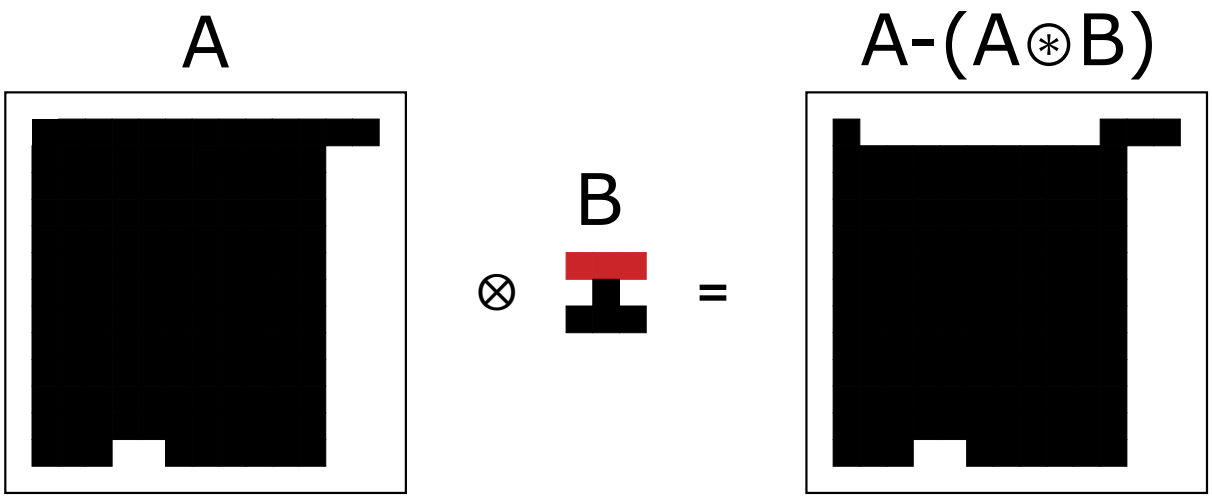



Figure 2.6: Image $A$ gets thinned by structure element $B$; $B$ consists of a hit part (black) and a miss part (red)

The optimal ratio of *hit* and *miss* structuring elements was analyzed in detail by [17]. The resulting array of sampling points left after binary morphological operations can then be used to estimate the best polynomial fit, using least-squares methods. Various structuring elements and the general comparison between the two methods, spatial voting and morphological processing, will be evaluated further in Part III.

## 2.3 Peak detection via correlation

A novel approach, for the detection of peaks of triangle-shaped surface profiles, i.e. windrows, consists of the utilization of correlation by a 1-D Gaussian kernel. This technique is especially useful for dense distance measurement maps, obtained by a stereo camera for example. Because of the density of the measurements, peaks of windrows are detected precisely with the correlation approach. This peak can then afterwards be used as a feature for gaining heading information. The method can be used for every row of an image of a stereo camera, where the shape of the windrow is visible. As an example for a single row, Figure 2.7 shows how the correlation with a Gaussian kernel function works.

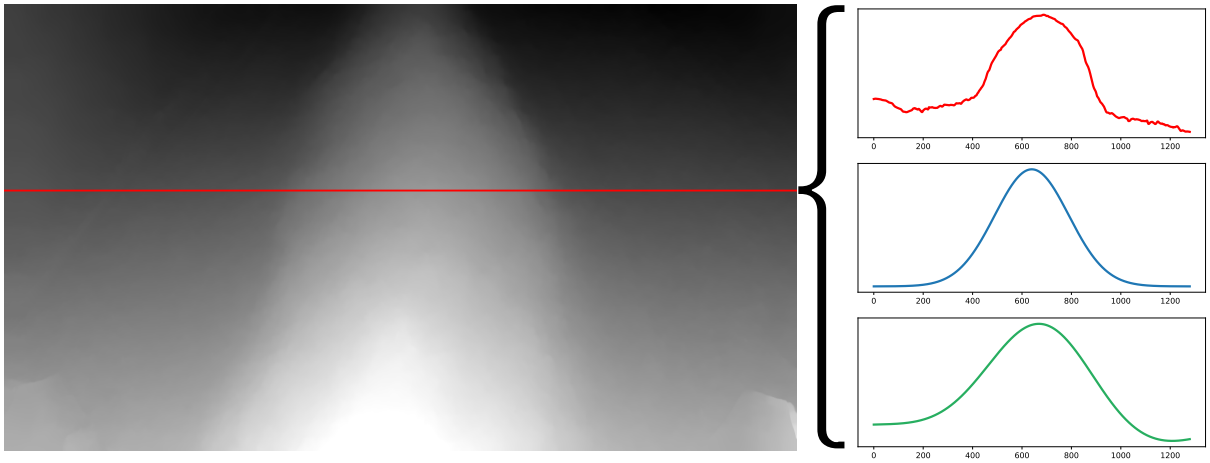



Figure 2.7: Schematic visualization of correlation with a 1D Gaussian kernel function; distance measurement(red), Gaussian kernel (blue), result of correlation (green)

One way to detect peaks in the data would be to get the highest value of the input data (red) directly. However, results of this method can be affected by erroneous depth readings or potential unwanted side peaks in the image. By correlating the input data with a kernel function (blue), similar to the shape of the windrow, before mentioned potential risks get neglected. The real peak of the triangle-shaped feature can be acquired by taking the highest value of the resulting function (green). This method can be applied for every row in the image. The resulting image coordinates (x,y) represent sampling points of the line derived from the triangle-shaped feature. The last remaining outliers in the data can be accounted for by utilizing robust methods like RANSAC. A potential result of this method is visualized in Figure 2.8

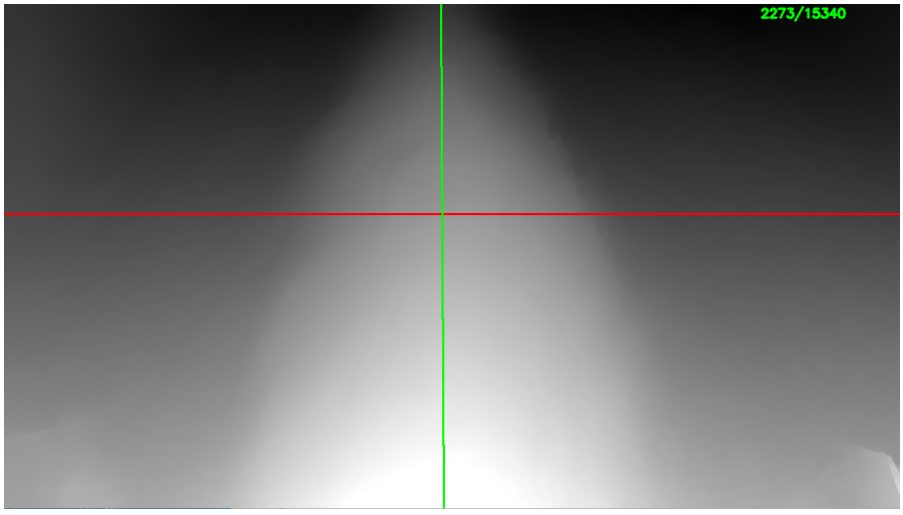



Figure 2.8: Line feature (green) derived from row-wise correlation method in combination with RANSAC

## 2.4 Visual odometry

The most advanced method for estimating relative heading, discussed in this thesis, is Visual Odometry (VO). This approach consists of the determination of relative translation and rotation derived from matched features between consecutive and/or concurrent images. It is also known as *pose estimation* or *extrinsic camera calibration* [18]. VO can be categorized in two groups, monocular and stereo vision VO.

**Monocular VO**

Monocular VO can be realized with low-cost cameras and deployed on readily available hardware. The processing power needed is low, in comparison to stereo vision system, because only one image has to be processed per time step. Therefore it is, up to now, the most common approach. The advantages of this method also come with some drawbacks. The most significant drawback is the scale uncertainty. By only using one image, the extraction of scale information is not straightforward and this can lead to a wrong estimation of the pose. This topic is discussed in more detail in [19].

**Stereo VO**

Stereo VO uses a binocular cameras with two lenses and separate image sensors. The greatest advantage of this system is, that depth information can be derived in a single time step by the method described in section 1.4. Due to a fixed and well-known stereo baseline, the scale of the image is retrieved completely from depth information. Furthermore, the greater number of features in both camera frames increase the tracking ability in consecutive frames. One major disadvantage of stereo cameras is the computing power and time needed to process the images and track the features. In addition to this, the calibration of a binocular stereo camera is more advanced and errors within this calibration have a direct impact on motion estimation [20].

Aforesaid approaches can be further divided by the method used to acquire motion information from subsequent frames. The most common method is called **feature-based** method, and extracts features from either concurrent or consecutive images, or both, and tries to track these. Various feature-based algorithms have been developed in the past decades, with ORB-SLAM [21] as one of the latest and more sophisticated. One major limitation of feature-based methods is, that only information which complies with the chosen feature type can be used. All other information in the image is discarded. **Direct methods** avoid this limitation by using pixel intensity in the image

as direct input for the algorithm. A state-of-the-art direct VO algorithm is LSD-SLAM [22]. This is a monocular VO algorithm and, in addition to highly accurate pose estimation, allows for creation of large-scale maps of the environment. Recently a mixture of the two, called **hybrid methods**, arose and became very popular. These methods, in combination with steadily growing computational power, combine the advantages of both, direct and feature-based, to get a very robust and highly accurate pose estimation [23], [24]. An overview of the evolution of Visual Odometry techniques can be found in [25].

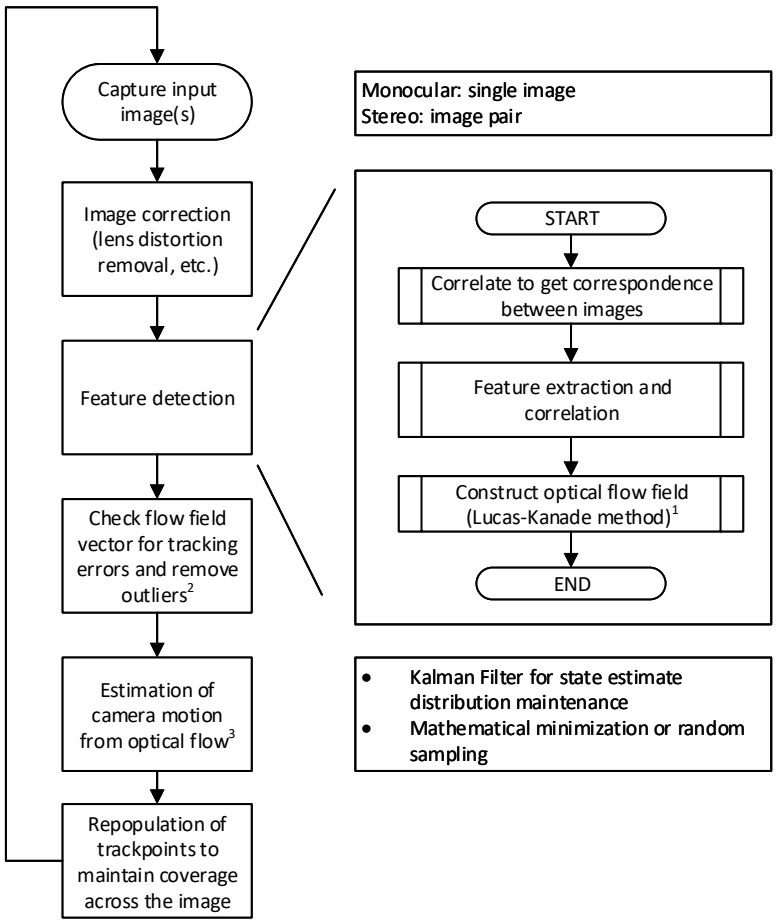



Figure 2.9: The basic workflow of a feature-based VO algorithm;
[1] differential method for optical flow estimation [26];
[2] flow field validation algorithm [27];
[3] pose estimation techniques [28], [29]

Figure 2.9 illustrates a simplified workflow of a typical feature-based VO algorithm. Parts of this thesis will deal with the investigation on how the feature-based VO approach compares to other methods described in chapter 2.

# 3 Inquiry of different sensor arrangements

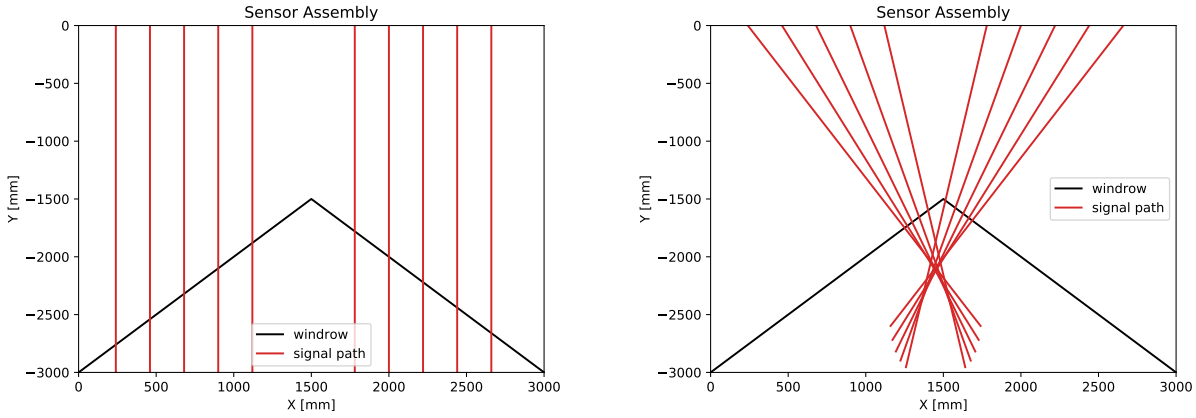
In order to acquire valid and robust measurements, the exact placement of the sensor with respect to the target has to be taken into account. Furthermore, the combination of multiple measurements strongly depends on the number of sensors used. In this chapter the placement of the different sensor types, discussed in subsection 1.1.1, will be analyzed.

## 3.1 Best placement regarding reflection angle

For active sensor technologies, such as TOF, the angle of the sensor regarding the target and therefore the reflection angle, plays a key role. The optimal scenario would be the direct reflection of a flat target. Here, all parts of the signal get reflected to the receiving unit of the sensor and we can derive the correct distance from the TOF of the signal. In practice, this is not always the case and one goal of this thesis is to get the distance from a sensor, mounted on the compost turning machine, to a windrow below it. In the optimal case it has the shape of an isosceles triangle. In the worst case, the windrow has many nooks and crevices, which reflect the signal diffusely.

Ultrasound, one of the TOF technologies used, depends on the reflection of sound waves. Even a reflection angle as high as 45° can be a problem for gaining reliable measurements. [30] Therefore, the arrangement angle of the sensor has to be adjusted in a way, that the signal path is approximately orthogonal to a windrow's surface. The term approximation is used, because in a real-life scenario, the perfect orthogonal signal path cannot be achieved in all situations. In Figure 3.1b, a sensor configuration for best orthogonal signal paths is shown. This example also considers the constraint of horizontal space available on the compost turning machine. One can clearly see, that the sensors have a nearly orthogonal signal path regarding the windrow's surface.

The major trade-off of this configuration is the spatial extent of the acquired surface points of the windrow. Meaning, almost all intersection points coincide with each other, forming two over-determined points on the windrow. In Figure 3.1a, the focus lies on the greatest spatial extent achievable. With this assembly, in theory, the best approximation of the form of the windrow can be acquired. Unfortunately, this is the worst configuration considering signal reflection. The best suited sensor assembly seems to be a compromise between the two extremes shown in Figure 3.1.




(a) best spatial extent

(b) best reflection angle of signal

Figure 3.1: Two examples of sensor configurations for surface profiling

## 3.2 Impact of sensor quantity

The quantity of sensors, and therefore measurements, does play a very important role when a surface profile has to be derived from range measurements. Generally, the more measurements there are, the better the resolution of the surface profile obtained. However, if the sensor placement is too dense, other problems may emerge.

For ultrasound ranging using low-cost sensors, one problem is the cross-talk between the sensors. Cross-talk occurs, because the sensors use the same modulation technique and therefore the signal is equal for each of them. Due to this, one cannot know if the signal emitted is also the received one. A possible solution to this problem is called sensor multiplexing. This topic will be treated further in chapter 5. Additionally, the computational power of the data acquisition module rises proportionally with the number of sensors. This leads to the usage of more expensive hardware. The problem of cross-talk for laser TOF based technology is not as critical as for ultrasound. The smaller beam-angle of the laser, in combination with more advanced evaluation

technology, yields a more robust system overall. Nonetheless, if laser sensors are packed very tightly, they are also prone to cross-talk and false reflections.

Aforesaid statements are only valid for active technologies, like TOF-based laser or ultrasound ranging. Passive technologies, like the stereo camera used in this thesis, cannot be a victim of cross-talk, even when using multiple sensors. Due to the wide Field of View (FOV) and the high resolution images used by the stereo camera, one sensor is sufficient to cover the whole area of interest. In the test-setup described in chapter 5, a compromise between resolution of the obtainable surface profile and economic factors has been made.

# Part II

# Practical setup and first investigations

# 4 Statistical analysis of sensor accuracy and precision

This chapter will deal with the statistical analysis of the distance measurements of each sensor type. Overall, three main statistical parameters namely accuracy, precision and robustness of the measurement are evaluated. The conditions evaluated are rising temperature and humidity levels and the reflection of various materials. The alignment of the sensor with respect to the target is another parameter in the laboratory tests.

## 4.1 Preliminary study of reflection and angle parameters

The utilized sensor technologies used, consist of two different measurement principles, *passive* and *active*. Active technologies have the disadvantage of needing a surface which has high reflective properties in order to work properly and in a robust manner. Especially ultrasonic distance measurements heavily depend on the composition of the target material. If the emitted acoustic waves approach a material which dampens sound, very little or no signal will be returned. On the other hand, if the material reflects the signal diffusely, the measurement is also deteriorated. The best target for active sensor technologies will be a surface with specular reflection, perpendicular to the signal path. The different reflection types are visualized in Figure 4.1.

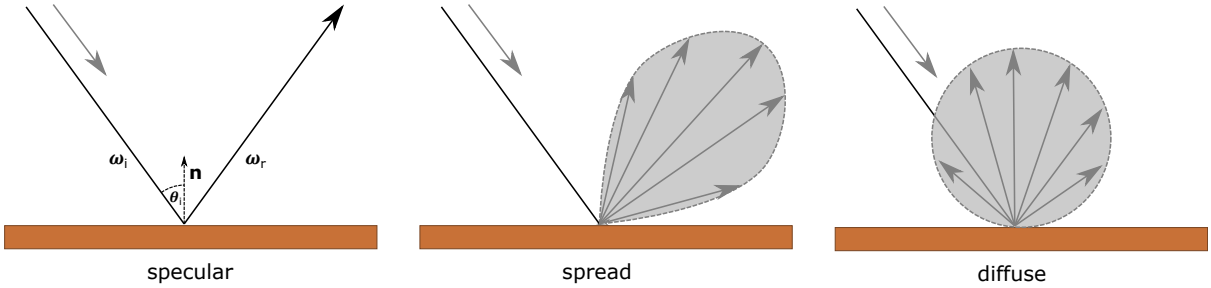



Figure 4.1: Types of surface reflections

For passive sensors, especially stereo cameras, the optimal surface for distance measurement is called a *Lambertian surface*. This is an ideal, diffusely reflecting surface which has the same radiance, independent of the viewing angle. Therefore, it's luminance is isotropic. In general, the reflection of a surface is described by it's Bidirectional Reflectance Distribution Function (BRDF)

$$f_{\mathrm{r}}\left(\omega_{\mathrm{i}}, \omega_{\mathrm{r}}\right) = \frac{\mathrm{d}L_{\mathrm{r}}\left(\omega_{\mathrm{r}}\right)}{\mathrm{d}E_{\mathrm{i}}\left(\omega_{\mathrm{i}}\right)} = \frac{\mathrm{d}L_{\mathrm{r}}\left(\omega_{\mathrm{r}}\right)}{L_{\mathrm{i}}\left(\omega_{\mathrm{i}}\right)\cos\theta_{\mathrm{i}}\mathrm{d}\omega_{\mathrm{i}}} \tag{4.1}$$

where $L$ is the radiance, $E$ is the irradiance, $\omega_i$ is the vector towards the light source, $\omega_r$ points towards the viewer and $\theta_i$ is the angle between $\omega_i$ and the surface normal $n$. The indices $i$ and $r$ indicate incident and reflected light.

All Lambertian surfaces comply with the Lambertian cosine law [31], and thus the BRDF value of such surfaces is constant. Therefore, the requirements for the optimal surface for active and passive sensors are reversed. In the next section, the three different sensors chosen in this thesis will be tested on surfaces with the reflective properties visualized in Figure 4.1. The angle between the sensor and the surface will also be altered during the tests. Altogether, every sensor will be tested with 9 different parameter configurations, yielding 27 independent data sets (Figure 4.2). These configurations will then be evaluated and compared regarding accuracy, precision and robustness.

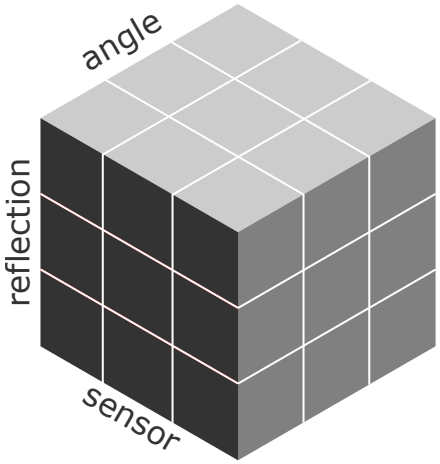



Figure 4.2: Visualization of measurement combination during testing

| Sensor | Reflection | Angle |
|---|---|---|
| Ultrasonic | Specular | 45° |
| Laser | Spread | 60° |
| Stereo Camera | Diffuse | 90° |

Table 4.1: Types of parameters used for data analysis

The parameter configurations, seen in Table 4.1, are visualized as combinations in Figure 4.2. Every sub-cube represents a specific parameter/sensor configuration. The execution of the tests will be further analyzed in the next section.

## 4.2 Accuracy and precision tests

In the process of testing accuracy, precision, as well as the distribution of the measurements, several materials with different reflective properties have been used. The materials and their reference distances can be found in Table 4.2. These distances have been measured and controlled by tape measure and laser DME. The laser used is a Hilti PD5 laser range meter [32] with a measuring range of 0.2 to 70 $m$ and an accuracy of $\pm 1.5\ mm + 20\ ppm$. The tape measure has EU accuracy class EG-II. Collectively, nine control groups have been created per sensor. Each control group consists of 1000 samples, recorded with 15 Hz update rate. The structure used for testing consists of an aluminium bar where the sensors have been mounted onto. Sensor centers have been aligned to ensure consistency regarding the measured distance.

| Material | Reflection Type | Reference Distance |
|---|---|---|
| Chalk Board | Spread | 1700 mm |
| Fine Fabric | Diffuse | 1658 mm |
| Mirror | Specular | 1725 mm |

Table 4.2: Materials used as test specimens for accuracy and precision tests

First and foremost, the data has been analyzed with regard to their distribution. Especially the check, if the data is normally distributed, played a key role in the analysis. Normal distribution of data is heavily desired in science, because on the one hand it is the most common distribution in nature, on the other hand statistical relationships become very tractable. In Figure 4.3, the sample quantiles and the theoretical quantiles from normal distribution ($\mathcal{N} \sim (0,1)$) are visualized in a qq-plot. The rows of the grid represent the reflection angles and the columns represent the three different reflection types tested. If the data is perfectly normal distributed it would fit the diagonal red line in the plot. The stereo camera (green) best fits the assumption of normally distributed data and only differs distinctly in the specular reflection case. Especially a combination of specular reflection with a reflection angle of 45° results in skewed data with a heavy lower tail. This behaviour supports the assumption, that highly specular reflective surfaces cannot be measured repeatably with a passive sensor like the stereo camera.

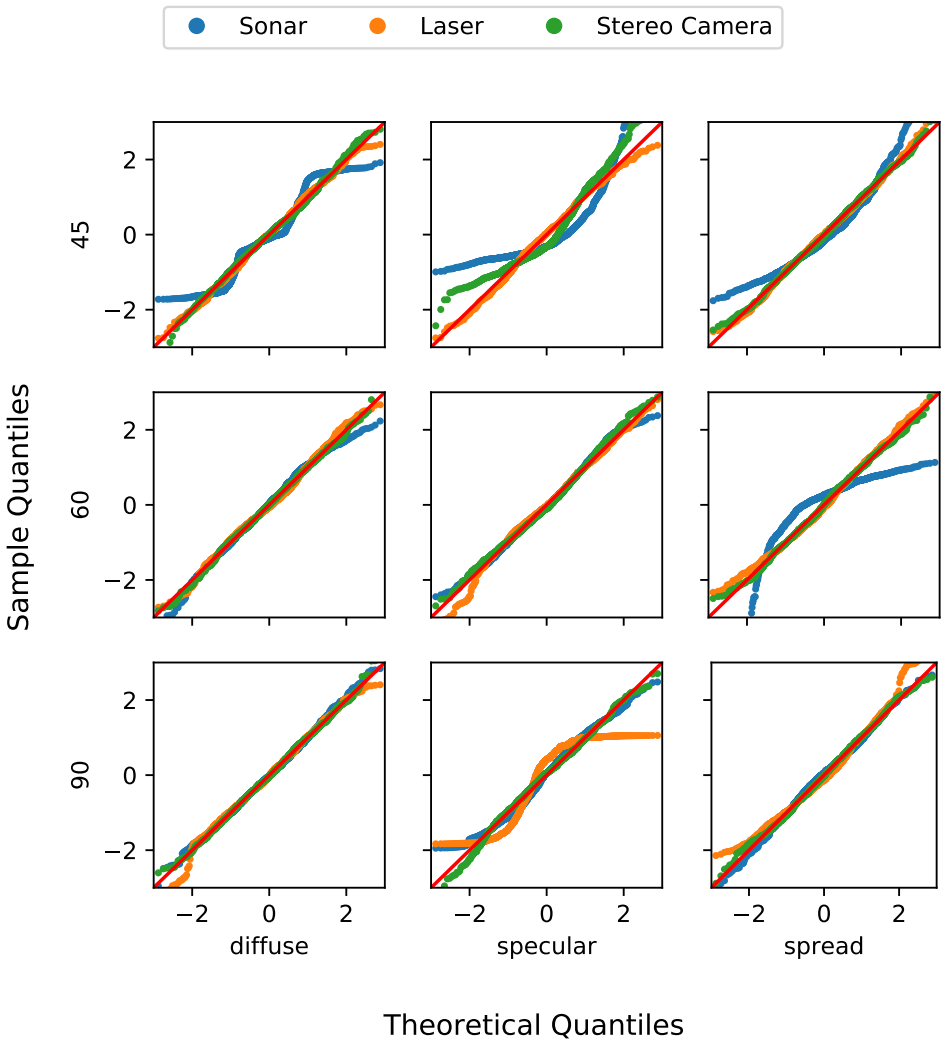
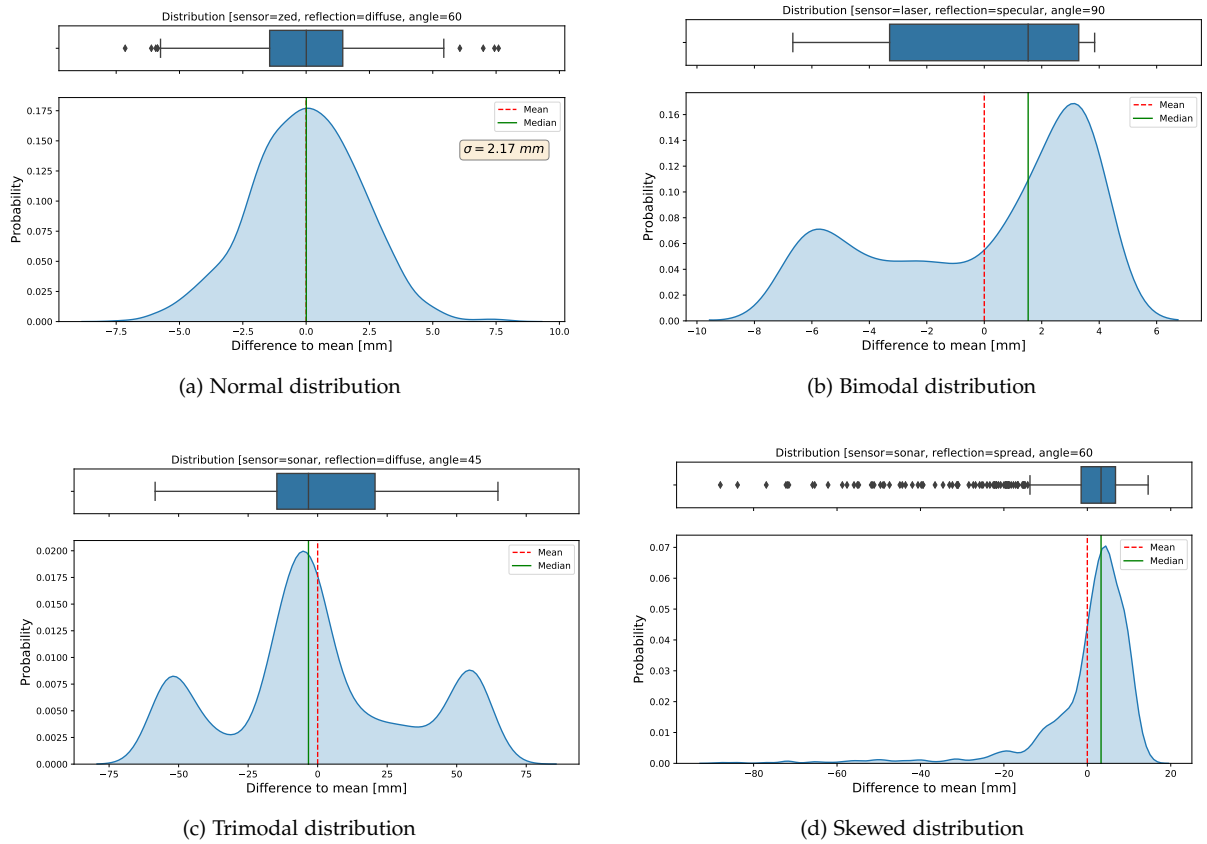



Figure 4.3: Quantile-Quantile plot the sample quantiles and the theoretical quantiles ($\mathcal{N} \sim (0,1)$)

The laser DME also had the most difficulties with repeatable measurements on specular surfaces. This may yield from the selection of a perfectly specular surface (mirror) in this example. The measurement principle of the chosen DME cannot cope with the given surface and results in a highly bimodal distribution. One of the peaks shows the real distance to the mirror and the other the distance to the mirror image. Otherwise the laser did compare favourably with the other sensors, with only minor tails in the upper and lower regions. The worst of the three, with regard to normally distributed measurements is the ultrasonic sensor. In the case of diffuse reflection with a reflection angle of 45° the distribution of the data has trimodal shape. As discussed in section 3.1, reflection angles greater than or equal to 45° cannot be measured accurately with the measurement principle of ultrasound. Highly skewed data in the spread scenario with angles less than 90° is also very prominent. The chalk board seems to spread the reflection of ultrasound very well when emitted with a non-orthogonal signal path.




(a) Normal distribution

(b) Bimodal distribution

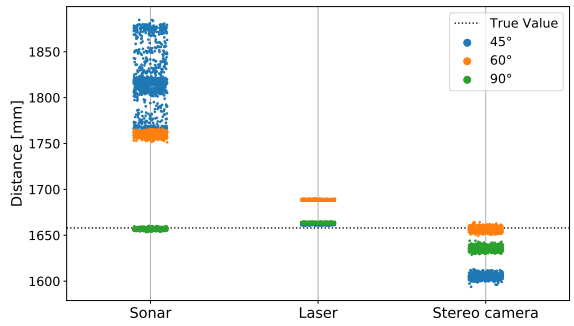
(c) Trimodal distribution

(d) Skewed distribution

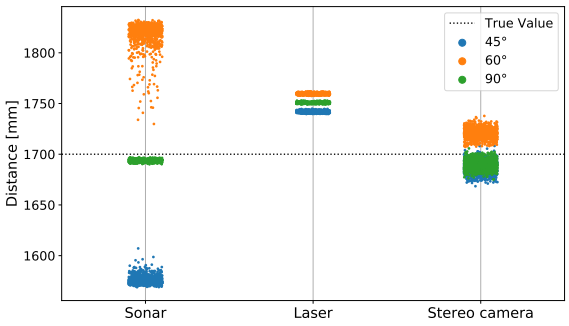
Figure 4.4: Comparison of extreme cases regarding distribution in recorded data

Figure 4.4 shows four different types of distributions. As a measure of variability, the standard deviation $\sigma$ and the Median Absolute Deviation (MAD) can be considered. Standard deviation is only a significant measure, if the data is normally distributed or close to normal distribution (e.g (a)). For bimodal, trimodal or skewed distributions

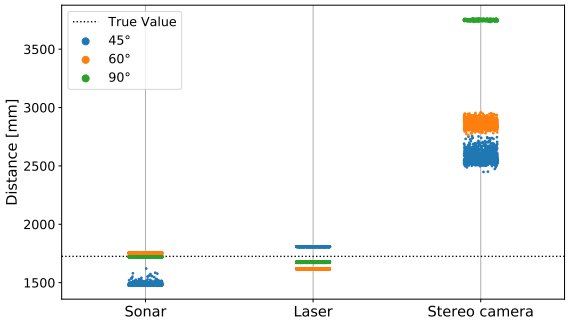
(b-d), the standard deviation cannot be used as a statistical parameter. The first and third quartile (visualized in the box plots) in combination with MAD better describe those distributions.




(a) Diffuse reflection




(b) Spread reflection




(c) Specular reflection

Figure 4.5: Comparison of extreme cases regarding distribution in recorded data

To get a better understanding of precision and accuracy of each sensor, Figure 4.5 shows all 27 cases described in section 4.1. The real distances, controlled by laser distance measurement and a tape measure, are shown as dotted black lines. Overall, it is clearly visible, that the laser sensor has the best precision of the three. For reflection angles less than 90°, measurements of the ultrasonic sensor tend to have the greatest distribution and thus the worst precision of all. Interestingly, when comparing samples with direct reflection paths (90°), ultrasonic measurements are the most precise and accurate. Therefore, these measurements have a very strong correlation with the chosen reflection angle. Laser measurements on the other hand, give the indication to be unaffected against reflection angle change. In the case of diffuse (a) and spread reflection (b) the stereo camera compares well with the results of the laser sensor. Specular reflection, as already seen in Figure 4.3, is a huge problem for passive sensors and therefore the accuracy of the stereo camera in this case is the worst of all. To summarize and visualize the context of accuracy of all sensors, Figure 4.6 visualizes the absolute difference between the mean of the measurements and the real values. This heatmap representation paints a very clear picture of the relations between reflection

type and reflection angle for each sensor. Two very distinct features also become clearly visible. First, the ultrasonic sensor has the best accuracy if a direct reflection path is given. The other prominent feature is the complete failure of the stereo camera when specular reflection is considered. Other than that, the stereo camera compares very well to the results of the laser sensor. When angle change is considered, laser and stereo camera show consistent results. Whereas the rate of accuracy of the ultrasonic sensor does drop rapidly with increasing angle.

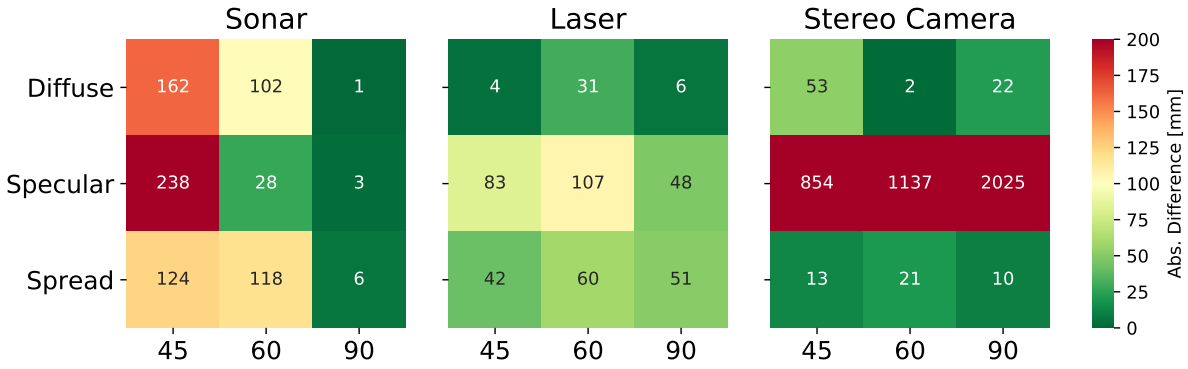



Figure 4.6: Absolute difference between the mean value of the measurements and the real value

# 5 Test setup

In this chapter the construction and composition of the test setup will be explained and illustrated. The utilization of ultrasonic and laser sensors in the low-price range required the development of a data acquisition module. The module handles the hardware communication to the sensors. The recorded information is saved to a SD card for further post-processing. The stereo camera is connected to a notebook, which records a side-by-side video of the two images. The whole setup consists of an aluminium profile, where the sensor components are mounted onto. The whole contraption is visualized in Figure 5.1.

Figure 5.1: Aluminium profile with mounted sensors; GNSS receivers (yellow) + IMU (orange) mounted on the sides for reference data generation; stereo camera and other DME mounted inbetween

## 5.1 Hardware design of data acquisition module

The hardware for the data acquisition module has been designed from the ground up. First and foremost the development of a Printed Circuit Board (PCB) (Figure 5.2), containing all hardware components, was necessary. The chosen communication protocol between the Microcontroller Unit (MCU) and the laser DME is Inter-Integrated Circuit (I2C). This is a bi-directional bus communication protocol, where each sensor gets an address to communicate with. The PCB also features a temperature and humidity sensor, used for compensation of the ultrasonic measurements. The ultrasonic sensors communicate with the MCU via normal input-output data lines. Additionally, the designed PCB features a voltage regulator to manage the power needs. The MCU used in this thesis is an ESP-32, developed by the company Espressif. This chip features a high clock speed, various communication protocols and in addition to this, WiFi and Bluetooth Low Energy (BLE).

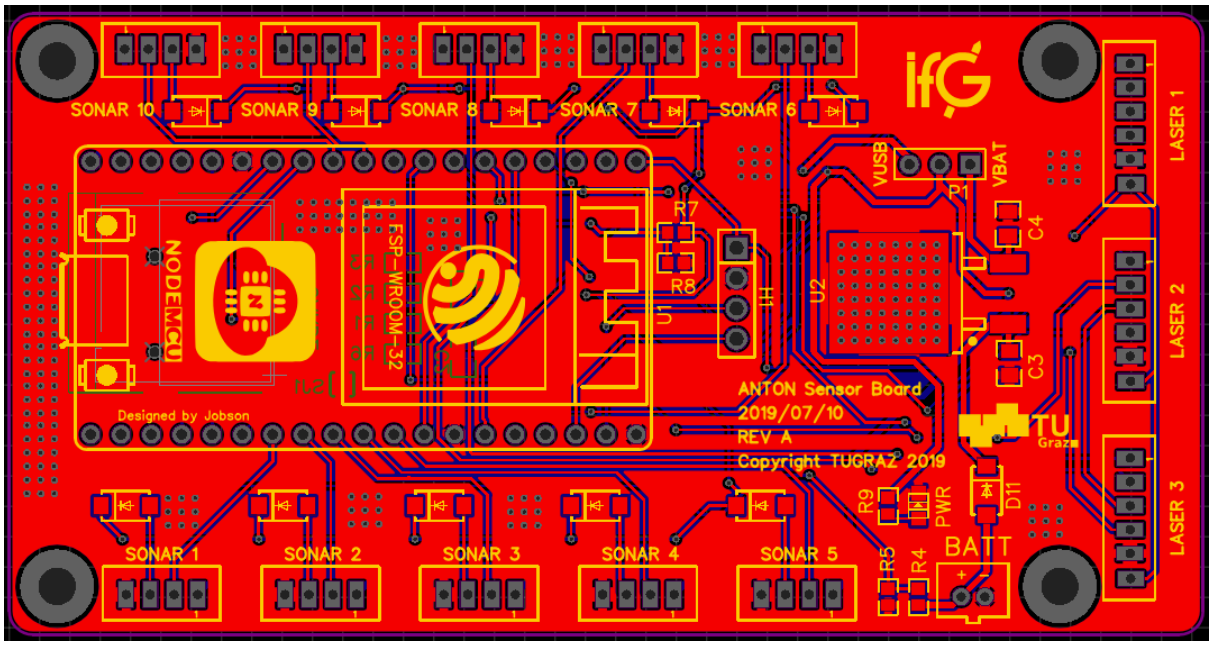



Figure 5.2: Hardware data acquisition PCB

### Multiplexing of ultrasonic DME

Each ultrasonic sensor features a trigger and an echo pin. These pins are necessary for sending and receiving the echo pulse of the ultrasonic sensor. The trigger pins are routed separately for each sensor to the MCU, but the echo pins share the same pin. This approach has been chosen, because of saving valuable I/O pins on the microcontroller and the fact that low-cost ultrasonic sensors have to be multiplexed to

use them in parallel. This extra-step must be executed, because all the sensors use the same modulation technique. Otherwise, this will lead to major cross-talk issues in data acquisition. Multiplexing is a technology where multiple signals are combined into one. In this application, a variant of multiplexing, called Time-Division Multiplexing (TDM), is used. This method splits the signal into individual frames, separated by a certain interval. As a result, only one sensor is activated at a given moment in time. Multiplexing ultrasonic sensors also has one major drawback, namely the update-rate of the measurement. Due to the speed of sound (0.34 km/s) being very slow, compared to the speed of light (300000 km/s) for example, the duration of sending and receiving a pulse is considerably longer. Summarized, this means the time between emitting pulses is longer and therefore it is slowing down data acquisition. As an example, imagine we want to multiplex 10 ultrasonic sensors. The range $d$ is about two meters. The used technology is based on two-way TOF, so the distance has to be doubled. This roughly yields an interval per sensor of

$$\Delta t = \frac{d \cdot 2}{\vartheta_{sound}} \approx 10\ ms. \tag{5.1}$$

Consequently, a complete cycle of the ten sensors is about 100 ms and we can see, that the update-rate of a single sensor drops linearly with the amount of sensors used.

## 5.2 Data-acquisition loop

The way the data is recorded, can be split up into two methods. The first and simpler method is the data acquisition of the stereo camera. Because the data of the test setup is analysed in post-processing, the only step here is to record the image frames with a notebook and save it on a hard drive. This has been done with a software provided by the stereo-camera manufacturer ZED Stereolabs©. The second way of collecting data is via the before mentioned hardware data acquisition module. The firmware running on this module executes two update loops in parallel. One for the ultrasonic range measurements and another for the I2C communication. The second loop consists of logging the measurements from the laser sensors and updating the environmental conditions every second to be able to correct the ultrasonic range measurements. The range measurements and their timestamps get logged to a SD card, mounted in the data acquisition module. In order to get the timestamp for the measurement and compare it to the stereo camera and reference data, the current Universal Time

Coordinated (UTC) is synchronized with a smartphone connected via BLE. A problem, which arises when working with different types of data and timestamps is, that there are many realizations of time available. Generally, a master time reference has to be chosen and other frames have to be converted to this one. In this thesis, this was especially important when comparing UTC to Global Positioning System Time (GPST) in chapter 8, where ground truth data was recorded with two GNSS receivers.

## 5.3 Time synchronization and time references

The importance of a fixed time reference for joining or comparing diverse data sets in science is immense. Therefore, one has to know which realization of a time reference is used and how to convert between those. At first, this section will elaborate some of the most important time references in data or computer science and their definitions. Afterwards it will be shown how the relations between the time references interact. An overview of various time references and their starting points is given in Figure 5.3

#### Universal Time Coordinated

Universal Time Coordinated (UTC) is the standard time frame used across the world. It is a combination of Universal Time (UT), which is dependent on earth's rotation and the Temps Atomique International (International Atomic Time) (TAI), an average time of 400 atomic clocks worldwide, defining the SI second. UTC combines the precision of atomic clocks in agreement with earth's rotation (UT) by introducing leap seconds, therefore it is *coordinated*. This happens when UT - TAI $>$ 0.9 seconds. It was officially introduced on the 1$^{st}$ January 1972 and is still used today as a reference for civilian time around the world.

#### Local Time

Local time is the utilization of UTC in a specific timezone. In Great Britain for example where GMT is the local time, it coincides perfectly with UTC because it was defined that way. Local time does also consider daylight-saving time.

### GPS Time

GPST is a continuous time scale, like TAI, and based on the atomic clocks of GPS monitoring stations as well as satellites. The starting point of GPST was January 6$^{th}$ 1980 at 00:00:00 (midnight). At this epoch, the difference between TAI and GPST was 19 seconds. GPST is kept synchronized with UTC at about 25 ns precision. Leap seconds do not influence GPST.

### UNIX Time

UNIX time is defined as the number of seconds elapsed since the UNIX epoch (January 1$^{st}$ 1970 at 00:00:00). Leap seconds are ignored and every day is treated as if it has 86400 seconds exactly. It is mostly used in operating systems for internal timekeeping and in computer science applications. The definitions state, that it is standard to UTC, but with the additional information that time prior to 1$^{st}$ January 1972 (start of UTC) is not defined precisely.

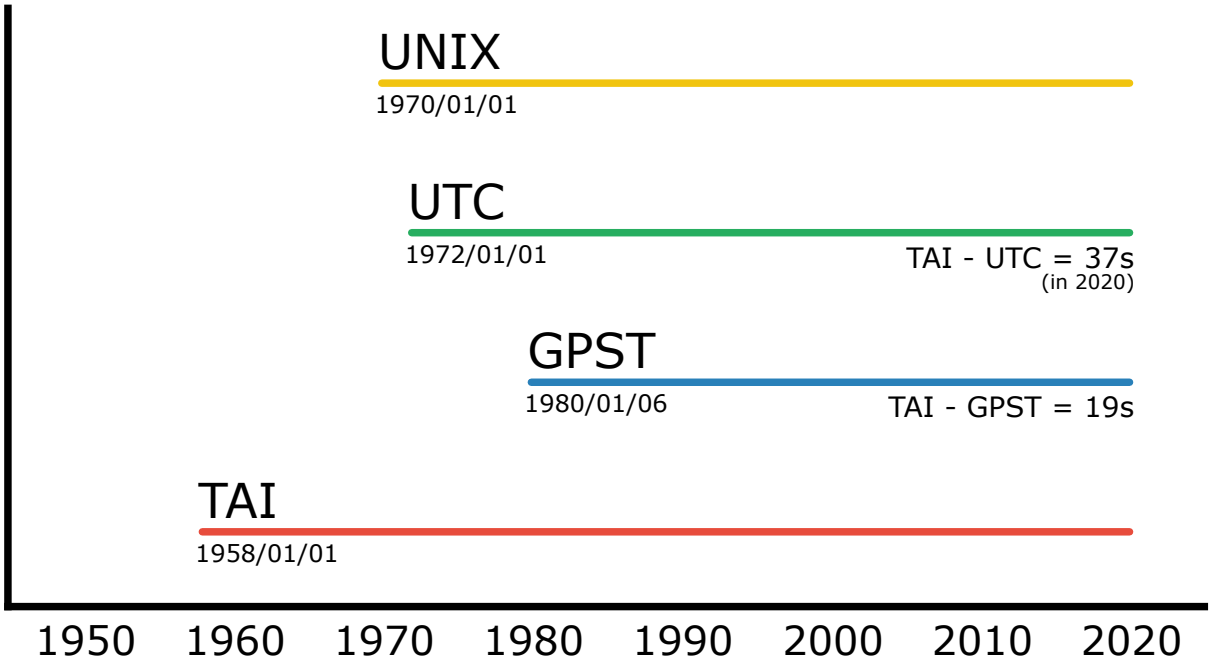



Figure 5.3: Four of the most important time references in comparison

As an example, if one wants to convert a GPS timestamp to UNIX time, the fixed time-offset has to be determined first (6.1.1980 UTC - 1.1.1970 UTC). Afterwards, the current value of GPS to UTC deviation has to be applied.

# 6 First measurement campaign

The first measurement campaign took place on the 20$^{th}$ August 2019 in Gnas, Austria, with hot and dry weather conditions. The chosen test site is a small private composting facility, run by a local farmer. The machine used on this site is a small and compact compost turner model A30 by the company Backhus. Some of the technical data of this machine are given in Table 6.1. During the first test run, a single windrow has been turned twice to get sufficient test data from all the sensors. The main goal of this first campaign was to get to know the dynamics of the machine and to check, if the acquisition module, in combination with the stereo camera, worked as intended. Furthermore, the two GNSS receivers and the IMU have been validated regarding data acquisition. The operator of the compost turner was also asked to steer the machine in wavy patters instead of the already straight shape of the windrow. This ensured that a wide range of distances is measured by each sensor. An overview of the complete test site, including machine, operator and windrow, can be seen in Figure 6.1.

| | |
|---|---|
| windrow width up to | 3 m |
| windrow height up to | 1.3 m |
| turning capacity up to | 700 $m^3/h$ |
| engine | 35.4 kw @ 3000 rpm |

Table 6.1: Technical data of Backhus A30 compost turner

Figure 6.1: Overview of test site in Gnas, Austria

## 6.1 First insights and analysis

This section will approach first insights regarding measurement setup, sensor configuration and the overall measurement procedure. First, the raw data of the ultrasonic measurements will be analyzed. In the next step, the recorded videos of the stereo camera will be examined more closely. As this was the first test of the prototype sensor configuration in combination with a newly developed hardware data acquisition module, there were some recording failures during data acquisition. Therefore, this section will only deal with time spans, where sufficient data has been collected. Unfortunately, the acquisition of data from the laser sensors was a complete failure during the first measurement campaign, with only one of the three sensors recording data infrequently. Thus, the recordings of these sensors will not be discussed in this section.

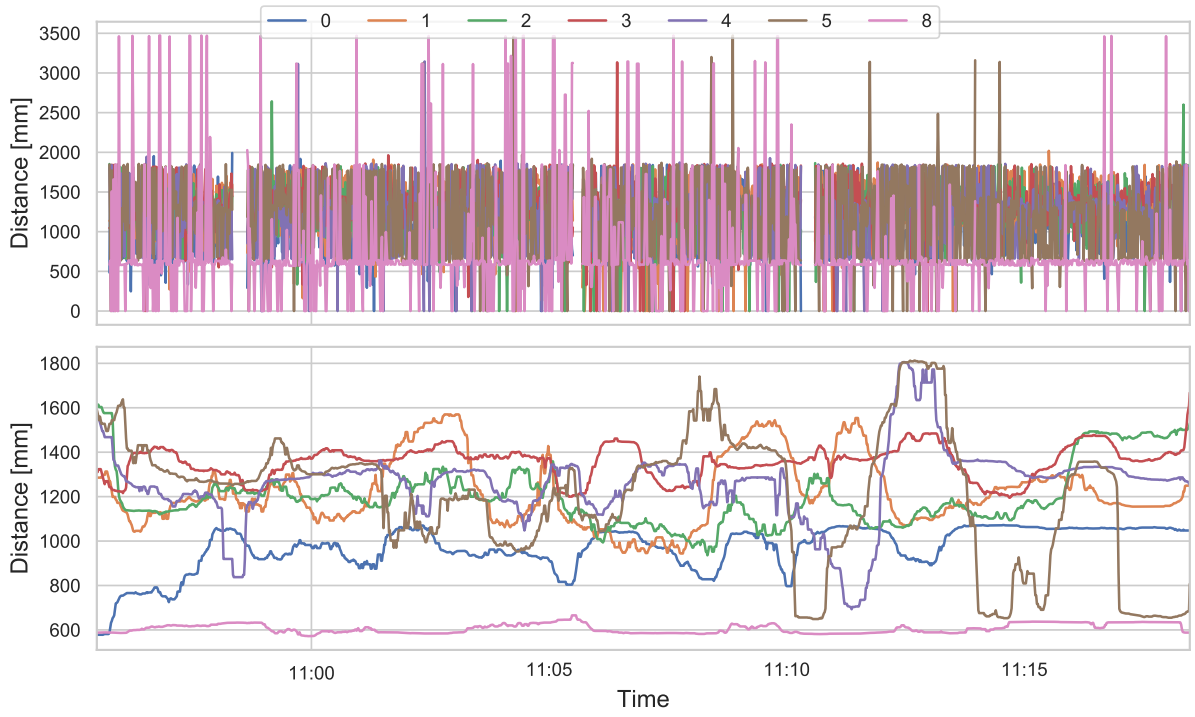



Figure 6.2: Upper figure: raw data of ultrasonic measurements, resampled to second intervals
Lower figure: data filtered with MMAD with a windows size of 60 seconds

The analysis of the ultrasonic measurements resulted in plausible data for 7 out of 10 sensors. Data of the other three sensors has been classified as implausible and will not be treated further in this section. The upper part of Figure 6.2 shows the recorded raw data of 7 ultrasonic sensors, resampled with an interval of one second. The whole time span of this dataset is about 20 minutes. It is clearly visible, that all the measurements are extremely noisy and in addition to that, some of the greatest deviations show up

repeatedly. The lower part of this figure shows the filtered data. A Moving Median Absolute Deviation (MMAD) filter with a window size of 60 seconds has been chosen. As the sensors are distributed equally left and right of the peak of the windrow, one could assume that the behavior of these two groups should be similar. To test this similarity assumption, a correlation matrix of the measurements has been computed. This matrix is visualized as a heatmap in Figure 6.3. Clearly, the auto correlation of each data set (main diagonal) results in a positive correlation factor with a value of 1, meaning equality. The assumption is, that the correlation factor calculated from measurements on the left and right side should be negative. For these datasets, this is just partly visible in the correlation matrix.

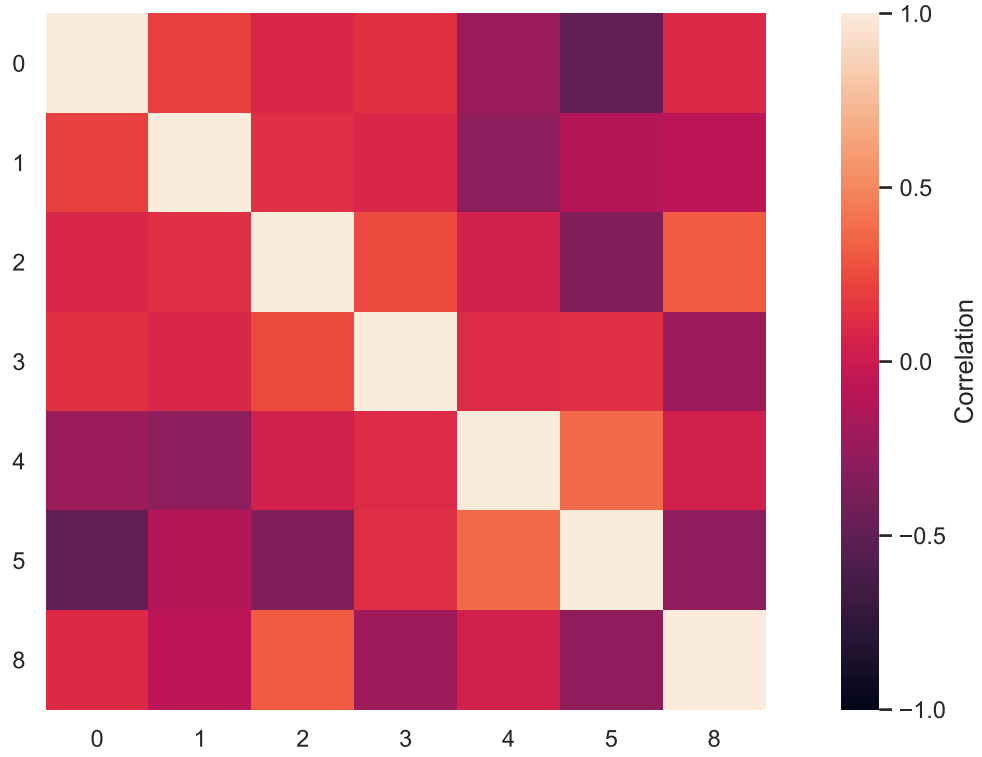



Figure 6.3: Correlation matrix of the ultrasonic sensor measurements

Data acquisition for the stereo camera was the most effortless of all three. The provided Software Development Kit (SDK) by the manufacturer Stereolabs© made it very easy to configure the needed parameters and start recording videos, which can then be analyzed in post-processing. The only problem was the size of the recorded video files. A single 20 minute drive through a windrow resulted in tens of gigabytes of generated material. This has to be taken into consideration when one wants to record multiple files. A single frame of a video file is visualized in Figure 6.4. This is a side-by-side, combined and rectified image of the left and right view of the camera. Every frame

is then later on analyzed and depth information can be extracted via the process described in section 1.4. The sequence of all images is also used as input for a VO algorithm.

Figure 6.4: Left and right view of ZED stereo camera by Stereolabs©

## 6.2 Summary of the first campaign

The first campaign was a very important step in understanding the system dynamics and evaluating the newly developed data acquisition module. The assumption for best arrangement of the hardware components was also inspected and analyzed for improvement. The assembly of the ultrasonic and laser sensors in particular was the most important factor to consider. This first test is also understood as a proof of concept regarding data recording and acquisition. In addition to that, data for later reference comparison has also been recorded via geodetic GNSS receivers in combination with a high-end IMUs. As described in the previous section, there have also been some complications, especially data acquisition for the laser sensors resulted in a total failure in the first test run. Some measurements did experience unusually high noise, when compared to the laboratory tests. Solutions for these problems will be tackled in the next section.

# 7 Improvements & lessons learned

This chapter will deal with ideas of improvement regarding data acquisition and general sensor configuration for a more successful second measurement campaign. The difficulties faced in the first measurement campaign will be used as input in this improvement analysis.

## 7.1 Measurement noise mitigation

One of the most prominent problems during the first campaign was the high noise examined in the measurements of the ultrasonic sensors. Recurring sudden spikes in the data have been another difficulty. A solution to these problems is the utilization of a signal filter with prior outlier detection. All the filtering techniques have been implemented in software rather than in hardware for easier adaptation and faster integration. As a signal filter, an implementation of the renowned Kalman-Filter has been used [33]. This algorithm works by recursively estimating the state of a dynamic process. It can also be seen as a Best Linear Unbiased Estimator (BLUE) for the state vector [34]. Values for the initial covariance have been taken from the results of chapter 4. The covariance matrix $\mathbf{R}$ consists of the variances $\sigma_S^2$ for sonar measurements and $\sigma_L^2$ for laser measurements. The bold font in this matrix should highlight, that all the factors are matrices containing $n$ variance values for $n$ sensors. Analogous to that, the state vector $\mathbf{x}$ also consists of vectors of distances $\boldsymbol{d_S}$ and $\boldsymbol{d_L}$.

$$\mathbf{x} = \begin{bmatrix} \boldsymbol{d_S} \\ \boldsymbol{d_L} \end{bmatrix} \qquad \mathbf{R} = \begin{bmatrix} \sigma_S^2 & \mathbf{0} \\ \mathbf{0} & \sigma_L^2 \end{bmatrix}$$

As an outlier detection scheme, simple thresholding has been implemented in software. Implausible distance measurements (i.e. measurements greater than five meters with a

total distance to the floor of three meters) are neglected and the previous measurement is used. This technique can be used, as all outliers detected in the first measurement campaign can be classified as extreme outliers.

## 7.2 Acquisition of ground truth data by surveying

The next improvement over the first campaign was to generate real ground truth data by surveying. This data can later on be used as a reference and an analogy to reality. This analogy can be made because of the high accuracy and precision of the total stations used in the process. Accuracy information for the used instruments are given in Table 7.1.

| | **MS60** | **TCRA1201** |
|---|---|---|
| **Angular Accuracy** | $1''$ $(0.3\ mgon)$ | $1''$ $(0.3\ mgon)$ |
| **Distance Accuracy** | $1\ mm \pm 1.5\ ppm$ | $2\ mm \pm 2\ ppm$ |

Table 7.1: Accuracy information of Leica total stations for measurements to a Leica 360°Mini Prism

The mini prisms have been mounted onto the left and right side of an aluminium profile. The distance between the two prisms is consistent with the distance between the two geodetic GNSS receivers. This means, that they share the same baseline, which simplifies the analysis and comparability in post-processing. The mounting positions of the prisms can be seen in Figure 7.1


Figure 7.1: Leica 360°Mini Prism mounted below geodetic GNSS receiver

Distance and angle measurements to the left and right prisms have been taken with a update-rate of 1 *Hz*. The prisms were tracked and locked with Automatic Target Recognition (ATR) technology by Leica. With the help of this technology, a motorized total station can automatically follow a prism locked beforehand and can take measurements at a given rate. The update rate of 1 *Hz* has been chosen because of the slow dynamics of the compost turner. Both instruments would have been able to take measurements with a frequency up to 10 *Hz*. To ensure time synchronization between measurements of the two instruments, both total stations were set up to track the same prism before the actual measurements started. The prism was moved up and down a few times to create a trajectory, which can then later on be used and compared in post processing to determine the time difference of the two instruments. During the process of synchronization, the measurement frequency has been increased to 10 *Hz* to get a better time resolution.

## 7.3 Adaption of the sensor array

In the first measurement campaign all active sensors (sonar and laser) have been aligned perpendicularly with their signal path pointing straight to the ground. This configuration has the best spatial extent possible. The downside of this method is that the reflection angle of the signal exceeds the value where reliable and robust measurements can be made. Therefore, the sensor array is adapted for the scenario of best reflection angle. Unfortunately, this method drastically decreases the spatial extent of the measurements. In the worst case, this reduces to two overdetermined points on the surface of the windrow. A comparison of both methods is visualized in Figure 3.1.

# 8 Second measurement campaign

This chapter will show how data acquisition and the measurement process in general has been improved during the second campaign. The analysis of the recorded ground truth data and first results will also play a key role in this chapter. Finally, the two campaigns will be compared.

## 8.1 Analysis of ground truth data

One of the most important aspects to consider, when combining measurements from two separate instruments, is the synchronization of time. Time synchronization can be split up into two categories.

### Relative synchronization

Relative synchronization tackles the coordination of the instruments with which the measurements have been taken. Figure 8.1 illustrates a trajectory, where one prism was tracked by two total stations at the same time. One can clearly see, that the shape of the curve is identical, but the time axis is shifted.

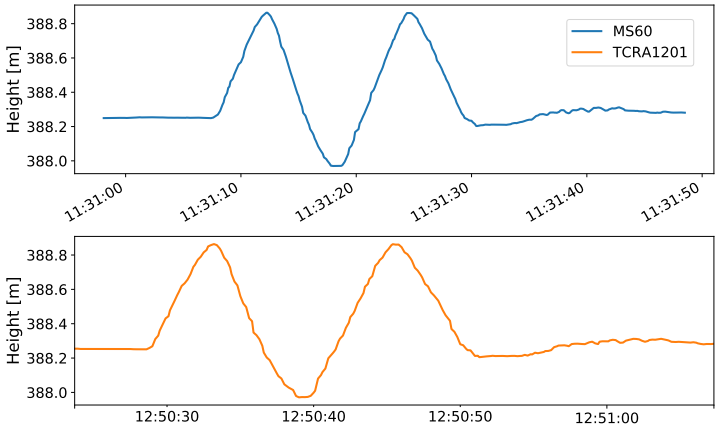



Figure 8.1: Height trajectory of two total stations with different time reference

To account for this shift in data, one of the two time references is fixed while the other one is shifted by the amount of

$$\Delta t_{\text{nolag}} = t_0^{\text{shift}} - t_0^{\text{fixed}}. \tag{8.1}$$

This sets the starting points of both datasets equal. If the recording of the data is started at the exact same time, this correction would be sufficient. In reality this is almost never the case. Therefore the residual lag in time has to be evaluated. This can be accomplished by cross correlation of the two datasets. The index, where the correlation coefficient reaches it's maximum, is equivalent to the lag, or time delay, in the units used. The mathematical definition of this calculation is

$$\tau_{\text{delay}} = \arg\max_{t \in \mathbb{R}}((f \star g)(t)) \tag{8.2}$$

with

$$(f \star g)(\tau) \triangleq \int_{-\infty}^{\infty} \overline{f(t-\tau)} g(t) dt, \tag{8.3}$$

where $f$ and $g$ are functions, $\overline{f(t)}$ is the complex conjugate of $f(t)$ and $\tau$ indicates the lag. The total amount of time delay $\Delta t$ between the two datasets can then easily be computed by subtracting the two terms calculated before.

$$\Delta t = \Delta t_{\text{nolag}} - \tau_{\text{delay}} \tag{8.4}$$

After this step, the two datasets are perfectly synchronized. A visual representation of the before mentioned steps is given in Figure 8.2. One important factor to mind is, that this time information is only relative. For example, although the shape of the calculated trajectory from values recorded by the two instruments is perfectly valid, the absolute time reference can be wrong to a large extent. So the last step is to transform the intrinsically correct data to a global time reference frame, to be able to compare the measurements to other frames. Once the data has been transformed into any time reference frame, it can easily be converted into any other time reference frame, as the transformation parameters between the frames are well known.

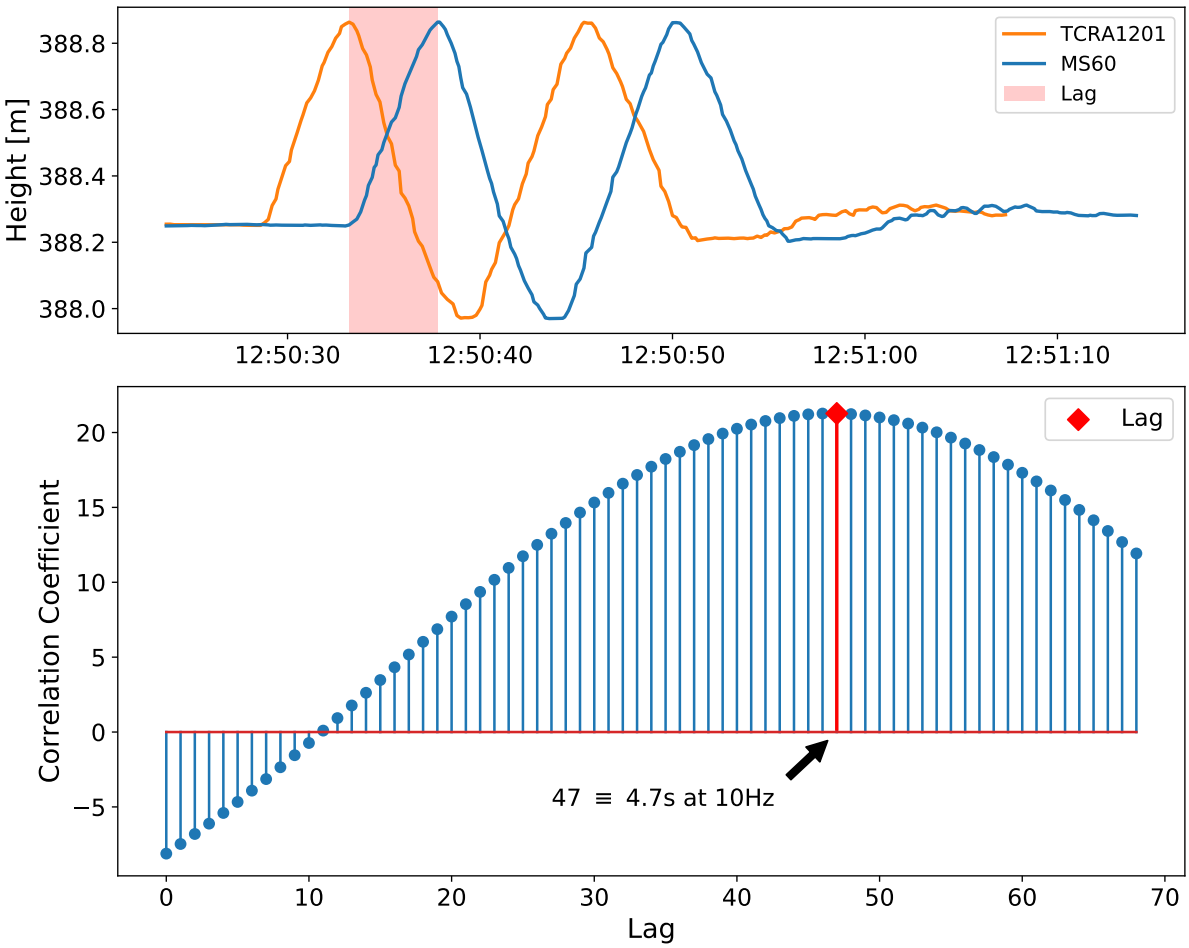



Figure 8.2: Lag $\tau_{\text{delay}}$ between two datasets; Time domain - upper figure; Correlation plot - lower figure

### Absolute synchronization

Absolute synchronization, in this context, is the coordination of the relatively synchronized data with a time reference like GPST, which can then later on be used to compare GNSS measurements with the generated ground truth data. The process is very similar to determining relative coordination. The trajectory created with geodetic GNSS receivers, via the moving baseline approach, is compared to the trajectory calculated from ground truth data. In addition to the computation of position, the moving baseline configuration allows for a simple technique to derive heading via

$$\gamma = \alpha + 90^\circ = \tan^{-1}\left(\frac{e_2 - e_1}{n_2 - n_1}\right) + 90^\circ. \tag{8.5}$$

Correlation between GNSS data and ground truth data has been calculated analogous to the procedure described in Equation 8.1 to Equation 8.4.

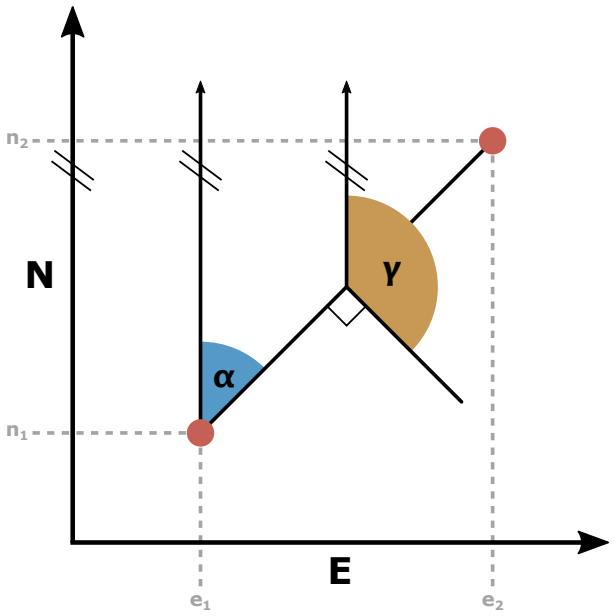



Figure 8.3: Visual representation of heading computation (Equation 8.5)

In the second measurement campaign, four windrows have been processed by the compost turning machine equipped with the sensors. The computed reference trajectories and their respective heading information, which will later on be used for comparison, are visualized in Figure 8.4.

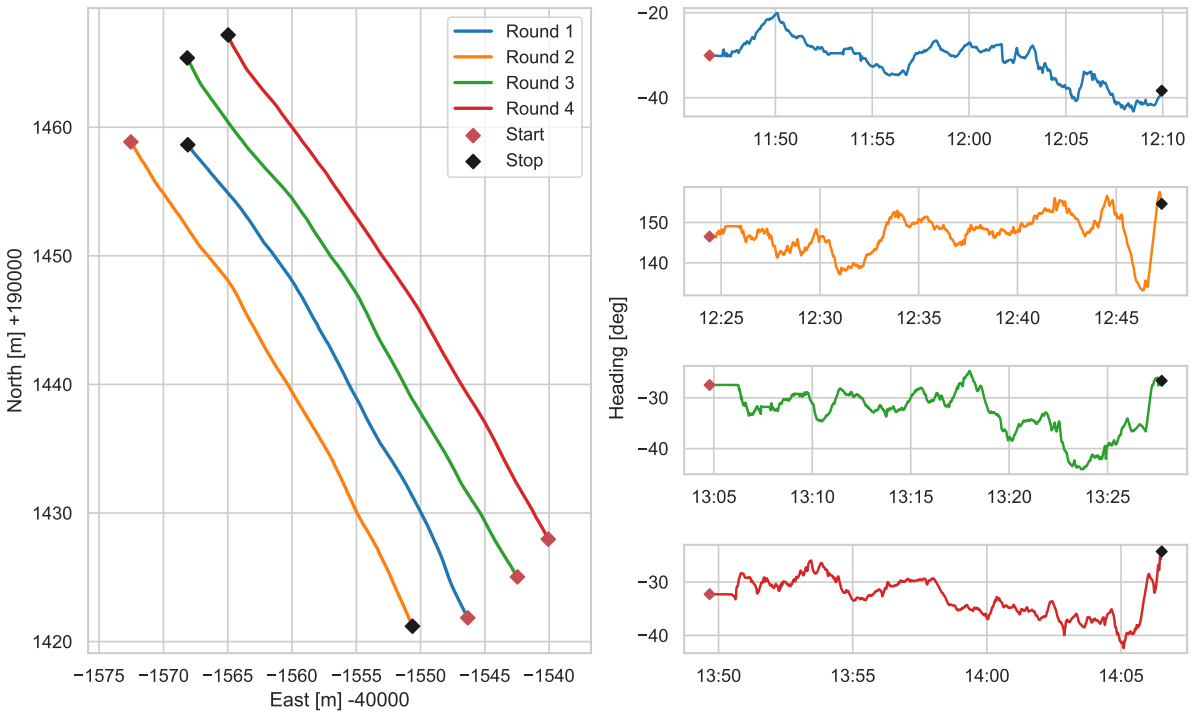



Figure 8.4: Reference trajectories computed from ground truth data

## 8.2 Analysis of recorded data

This section will deal with the analysis of the recorded data during the second measurement campaign. The procedure of data collection is similar to the first campaign, with the distinction of modifications elaborated in chapter 7. All sensor technologies in the second test run provide viable data for analysis. First, ultrasonic data will be examined and compared to the first test, second laser data will be investigated for the first time and lastly some modifications of the parameters for recording video with the stereo camera will be discussed.

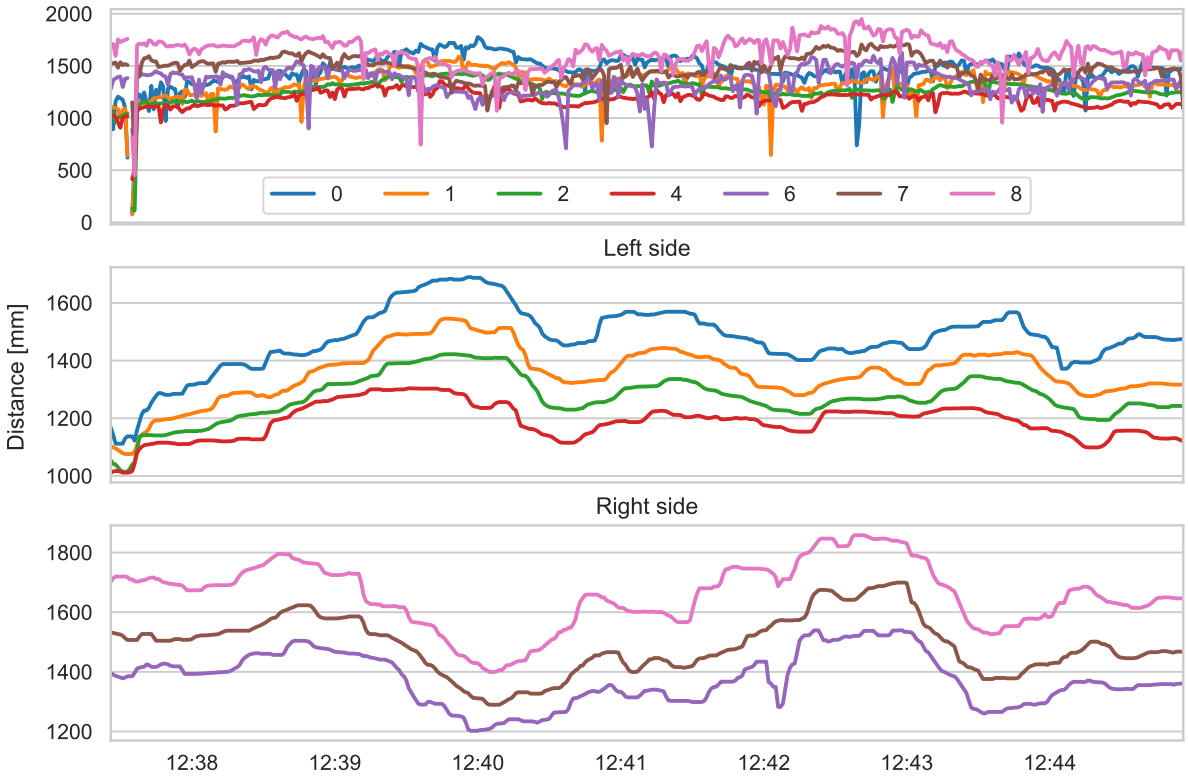



Figure 8.5: First figure: snippet of raw data of second round, resampled to intervals of one second
Second/Third figure: left and right side of windrow data, filtered with MMAD with a window size of 60 seconds

The second test run resulted in viable data for 7 out of 10 ultrasonic sensors. This number is equal to the number of working sensors in the first campaign. The only difference is, that the distribution between left and right side is more consistent. When comparing Figure 6.2 and Figure 8.5, the first thing clearly visible is, that the mitigation methods described in chapter 7 helped a lot in decreasing measurement noise. Extreme outliers, distinctive in campaign one, were also decreased with the help of setting a threshold for implausible values. The separation of measurements between left and right side of the windrow indicates, that the similarity between

those is more prominent in the second campaign. The comparison of the groups (left and right side) is closer to the assumption, that the sensor outputs of those groups should be negatively correlated. This assumption can be validated with the help of generating a correlation matrix of the measurements, illustrated in Figure 8.6. The correlation values confirm a strong positive correlation between the measurements on the same side and negative correlation for measurements of opposite side. Overall, ultrasonic measurements recorded during the second campaign promise to be of higher quality for further processing. The amount of data collected during the second run is also greater and therefore analysis of this data will be more expressive. All the improvements made during the second campaign seem to have a positive impact on measurement quality and robustness.

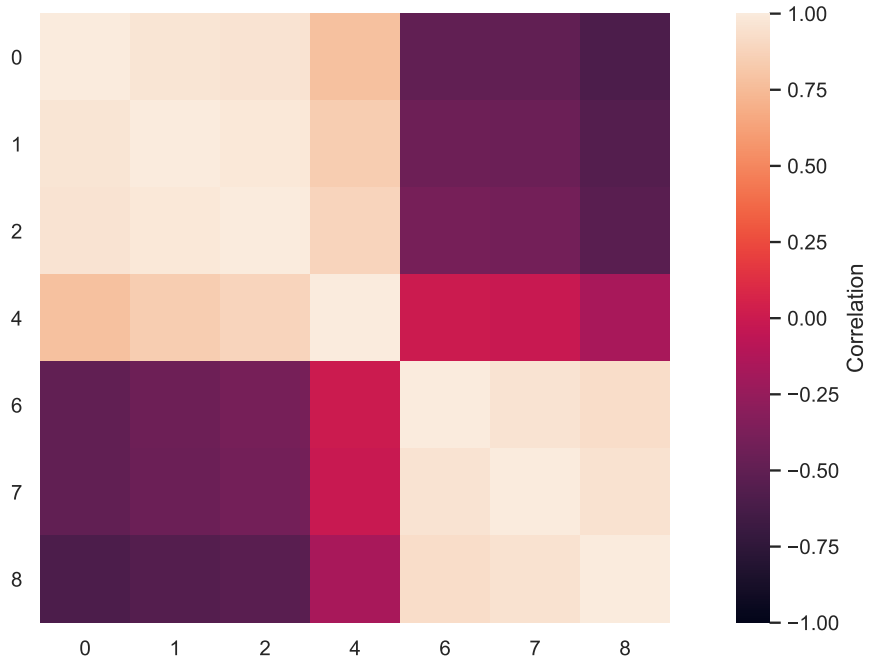



Figure 8.6: Correlation matrix of the ultrasonic sensor measurements in round two

During the second measurement campaign, all three laser sensors recorded data in every lap. Despite all the improvements made in comparison to the first campaign, data from the laser sensors is still very noisy and does not show negative correlation between measurements on opposing sides of the windrow (Figure 8.7). Considering the excellent results of the laser sensor on various types of materials with different reflective properties during the laboratory tests (section 4.2), this outcome was unexpected. One explanation for this behaviour could be, that particle dust hovering in the air, in the sensor's view disturbs the laser more than the ultrasonic sensors. A very important property to consider, when dealing with dust particles suspended in the air,

is the wavelength of the sensor. In such a scenario, sonar technology has an advantage over laser because of the larger wavelength used. Thus, resilience against particles in the sensor's FOV is strongly correlated with the measurement principle. This problem has also been discussed by [11].

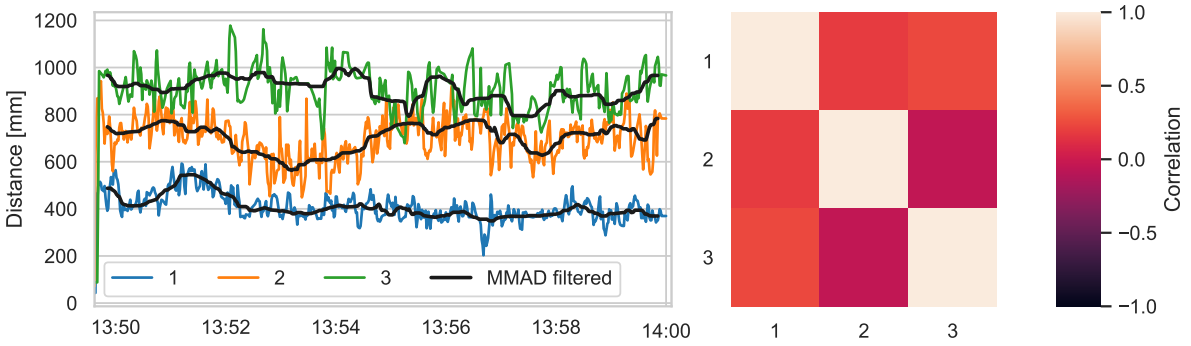



Figure 8.7: First figure: raw data of fourth round, resampled to second interval and overlayed with MMAD filtered data
Second figure: Correlation matrix of the laser measurements

The process of data recording for the stereo camera has not changed in comparison the the first measurement campaign. The only additional factor considered, was that enough space on hard drives was available to ensure data collection for every round.

## 8.3 Summary of second campaign

To summarize this chapter, it has been validated, that the improvements made after the first measurement campaign led to more reliable and robust measurements overall. Especially the reduction of noise in the ultrasonic measurements, combined with the mitigation of extreme outliers resulted in data of superior quality in the second campaign. Laser sensors, unfortunately, did not perform as expected. The quality of these measurements, for the major part, was very poor and noisy, compared to ultrasonic measurements. Because of this result, is has been decided against acquiring more of these sensors and undertaking a third campaign. The idea was to test a small quantity of laser sensors at first, with the outlook to extent the amount, if this type of sensor proves to be suited for distance measurements on windrows. As a consequence of low quality and low availability of laser data, chapter 9 will only deal with showing results of the algorithms described in chapter 2, applied to ultrasonic and stereo camera data.

# Part III

# Final analysis and conclusion

# 9 Results

This chapter will show the results of this thesis and try to compare them regarding quality, robustness and accuracy. Data used as input for this chapter was recorded in Gnas (small village in Styria, Austria) on the 28$^{th}$ October 2019. The weather conditions were cloudy with scattered rains at the end of the measurements.

**Preparations**

First and foremost, results from ultrasonic distance measurements will be evaluated and analyzed. Data acquisition has been successful in two of the four laps, with sufficient distance measurements on the left and right side of the windrow. The following section will deal with the comparison of the approaches shown in chapter 2. As input data, distance measurements from lap two and lap four will be taken. Approaches, listed under section 2.2 *Surface Profiling*, need an arcmap as input. The arcs for this arcmap have been generated using information about measured distance, arrangement angle and the FOV of the ultrasonic sensor. Figure 9.1 shows an example of the aforesaid arcmap at a specific epoch in time. The red rectangles describe the position of the ultrasonic sensors on the aluminium profile used during the measurement campaigns. Blue dots indicate sampling points of an arc computed from the parameters distance and arrangement angle of one sensor. The arc segment, consisting of many sampling points, stretches over the whole FOV of the sensor. The time span for measurements for each epoch in an arcmap has been 4 seconds. This decreases the temporal resolution of the computed heading information, but ensures to have sufficient arcs in the arcmap for further processing an thinning.

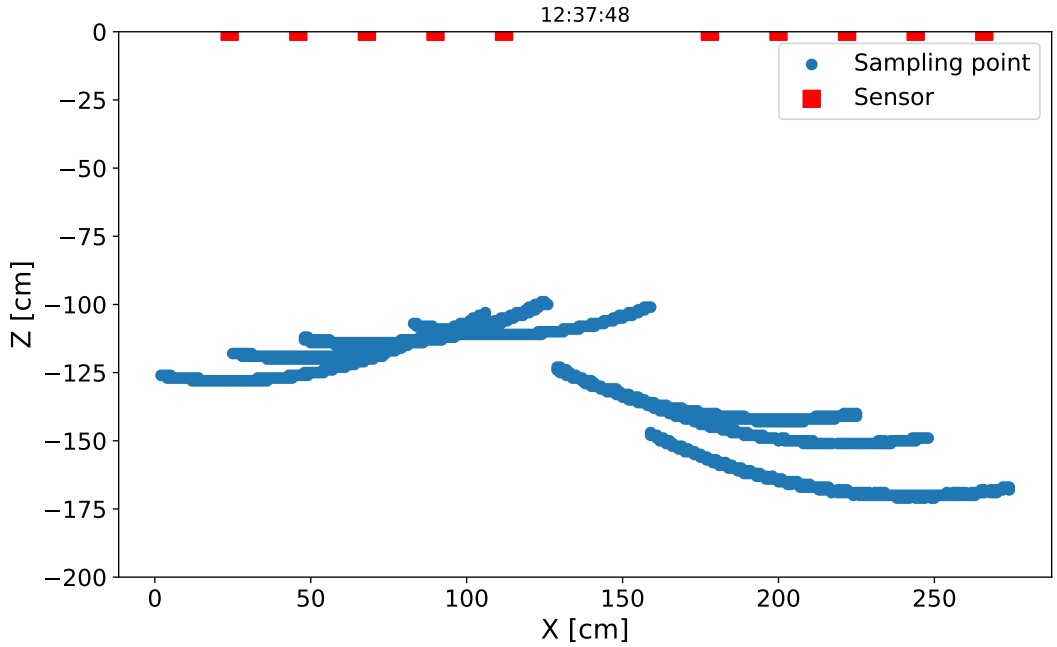



Figure 9.1: Arcmap of the profile of a single measurement epoch

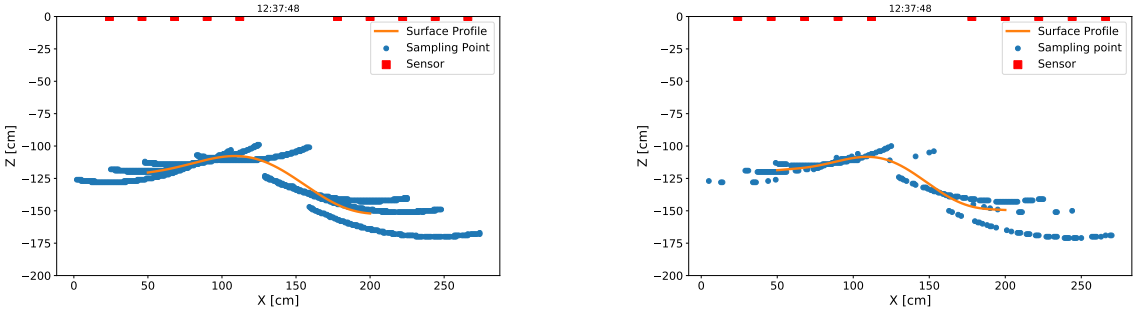
Two methods for arcmap thinning have been analyzed, namely *Spatial Voting* and *Binary Morphism*. For spatial voting, the right threshold value has to be found for the technique in order to work properly. Binary morphism on the other hand depends on the kernel chosen for the morphological operations. The parameters used in both cases can be found in Table 9.1.

| | |
|---|---|
| **Spatial Voting Threshold** | 15 |
| **Binary Morphism Kernel** | $3cm\ x\ 3cm$ |

Table 9.1: Parameters for data thinning methods *spatial voting* and *binary morphism*

Both methods have been applied to a 200 *cm* by 300 *cm* matrix with 1 *cm* resolution. The idea is to remove sampling points which are separate from clusters of sampling points generated by the real surface profile. For this to work properly, a large amount of overlapping arcs with various orientations have to be present in the data. If those arcs overlap in regions, where the resulting shape resembles the true surface profile, the outcome can be called a success. [14] has proven, that this technique works very well with a wide range of sensors. The attempt in this thesis, was to use this approach in every epoch of data recording and therefore derive the shape of the windrow below the compost turner. However, the low availability of reliable measurements from all sensors, paired with the fixed angle during a ride alongside the windrow, resulted in

sparse arc maps which do not replicate the true shape of the windrow very well. Both techniques could not be used to their fullest ability, because the sampling points in those sparse maps would vanish if the parameters in Table 9.1 are set too confidently. Nevertheless, peaks of the windrows have also been estimated from arcmaps with both thinning methods, to compare them against the simple approach section 2.1. Figure 9.2 visualizes the results of the two thinning strategies used.




(a) arcmap thinned by a spatial voting scheme

(b) arcmap thinned by binary morphism operations

Figure 9.2: Two arcmap thinning strategies

The results of spatial voting (Figure 9.2a) are almost identical with the unprocessed arcmap, because the distance information of all sensors has not changed much during the 4 second interval. Therefore, the time period between epochs would have to be even longer in order to thin out data with a spatial voting scheme. It has been decided against prolonging the interval, because it diminishes the temporal resolution of the method. Binary operations with a 3x3 kernel show a more thinned-out arcmap, but it is also far from optimal. Larger kernel sizes led to very sparse arcmaps, questioning the sense of using arcmaps in the process altogether.

The next sections of the results will deal with the comparison of the used methods. At first the results of the ultrasonic sensor array will be presented. Afterwards, methods using the stereo camera data as input will be compared against each other. Lastly, the best of all methods will be highlighted and compared in more detail to results from ground truth measurements. All heading plots in chapter 9 show a comparison of heading computed from ground truth data and heading computed from algorithms, described in chapter 2. To be able to compare them in an absolute manner, the start values have been equalized by initializing them with the value taken from ground truth data. As a measure of resemblance, the Root Mean Square Error (RMSE) is given for all comparisons.

## 9.1 Ultrasonic sensor

As shown in section 8.2, ultrasonic sensors did only collect sufficient data for a comparison in the second and fourth lap. In-depth analysis, regarding the two methods for heading determination from range measurements for ultrasonic sensors (chapter 2) resulted in failure when long-term stability is considered. The two methods described before, *spatial voting* and *binary morphism*, were not able to create a dataset with sufficient temporal solution which could be taken for comparison. Therefore, only results from the *simple approach* will be visualized and compared against methods using the stereo camera as sensor. One aspect to note is the temporal extent of the shown results in the next sections. For ultrasonic sensors it was only possible to get comparable results on a much shorter time frame.

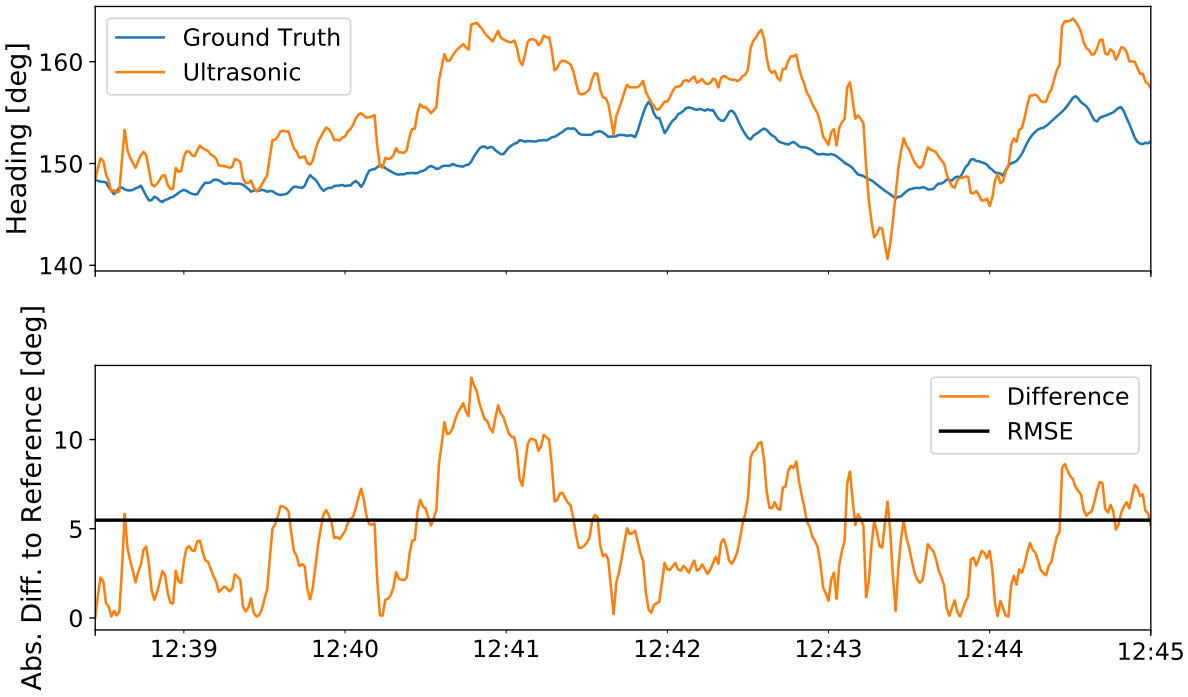



Figure 9.3: Comparison of ultrasonic results to ground truth data in second lap

When looking at Figure 9.3 and Figure 9.4, it is obvious that even over short periods of time, the results of ultrasonic measurements compare not very well to the reference. The RMSE is about 5 degrees in both laps. During these 5-10 min periods, sudden spikes with heading differences up tp 10 degrees are also present. These inaccuracies over the course of several minutes add up to a difference in heading, with respect to the reference, which is simply not practical for autonomous machines. Considering the high RMSE at a temporal extent of only 5-10 minutes, it becomes clear, that using

ultrasonic measurements to determine the heading of a compost turner with respect to a windrow is not feasible.

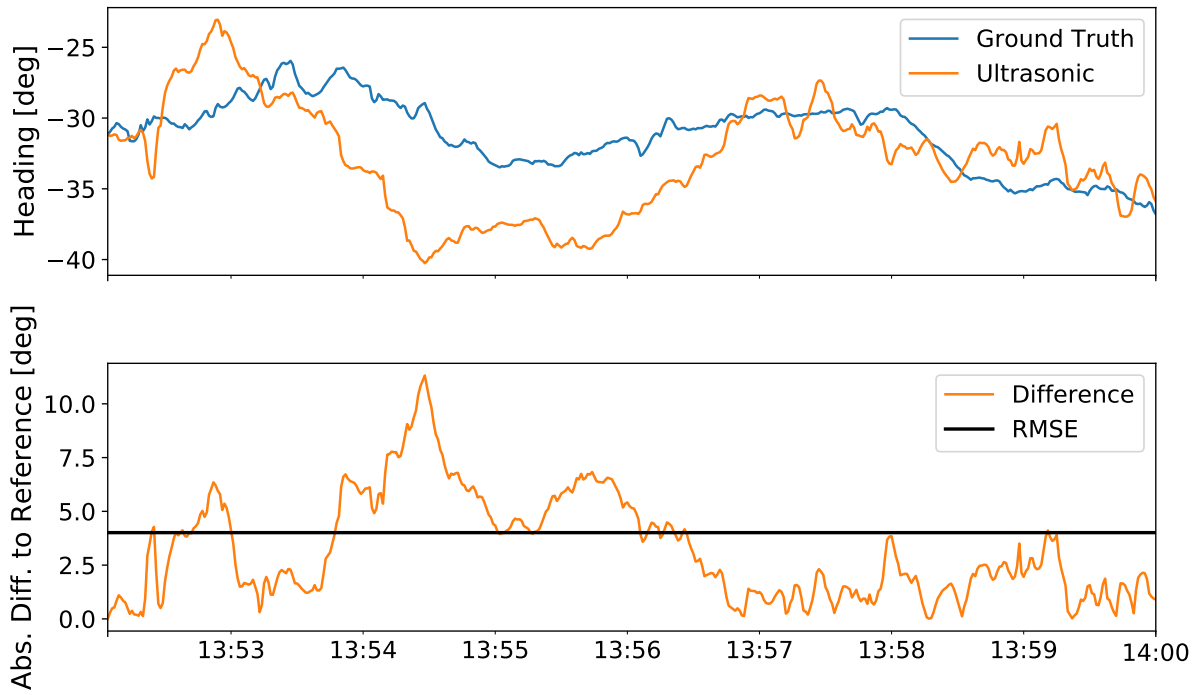



Figure 9.4: Comparison of ultrasonic results to ground truth data in fourth lap

To summarize, it has not been possible to achieve a robust and reliable method for heading determination by only using ultrasonic sensors. Improvements could be made by using more sensors overall, or by arranging them in a different manner to get better results from spatial voting or binary morphism. The best way to improve these methods would be, to create an array of moving sensors, in order to acquire a diverse set of distance-angle combinations for generating arcmaps. More reliable data acquisition in combination with more sophisticated noise mitigation methods are also ways to improve the system overall.

## 9.2 Stereo camera

This section covers the results gained from stereo camera image data. The first method, called *correlation method*, only uses depth images provided by the camera to derive heading information. The second method, *visual odometry*, uses both images in combination with highly sophisticated algorithms to derive very precise attitude or pose, including heading information.

The correlation method heavily depends on the correct detection of the windrow's peak. Therefore, if the shape of a windrow is very irregular or far from the shape of a triangle, this method has major problems in deriving heading information. This behaviour can clearly be seen in Figure 9.5. During the second lap, the compost turner drove alongside a freshly elevated windrow, where huge chunks of manure distort the triangular shape of the compost. Therefore, the location of the peak changes rapidly, yielding a trajectory heading which does not resemble the reference at all. In such a case the correlation method fails miserably.

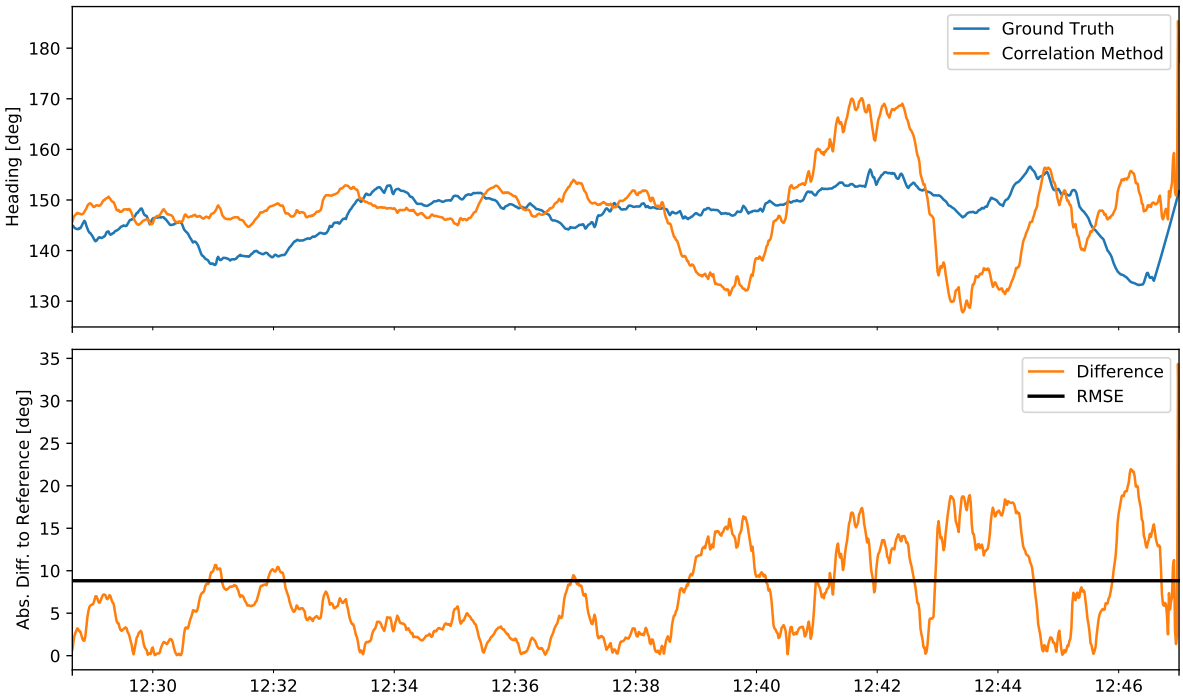



Figure 9.5: Comparison of results from correlation method to ground truth data in second lap

Figure 9.6 on the other hand, shows the result of image data acquisition of an older and triangle-shaped windrow. Here, the correlation algorithm can deliver reasonable results with a RMSE of about 4 degrees over the whole lap. During the end of the lap, the two datasets (correlation and reference) start to drift apart, resulting in differences of up to 10 degrees. The improvements over the results from ultrasonic sensors do not seem to be much, but the results of the correlation method cover a wider timespan over the course of a complete lap. In the case of ultrasonic sensors, the sensor array was not able to reliably record data for a complete lap and therefore only short periods could be examined. Although the correlation method, using depth images, seems to be an improvement over ultrasonic heading determination, the reliability of this method, considering less optimal windrow shapes, is very poor. Therefore, this method is also

not a worthy candidate to derive heading information from distance measurements. As referenced in chapter 1, the first requirements were the following:

- Reliable measurements regarding the material compost;
- Accuracy in the range of centimeters;
- Undisturbed by environmental conditions (temperature, humidity);
- Mid-to-low price range per sensor;
- Dust particles should have little to no influence on the measurement.

These requirements can be extended by

- Error of derived heading in comparison to the reference (RMSE) should be less than 1 degree over the course of a complete lap.

This will ensure that the compost turner is able to follow a windrow very precisely.

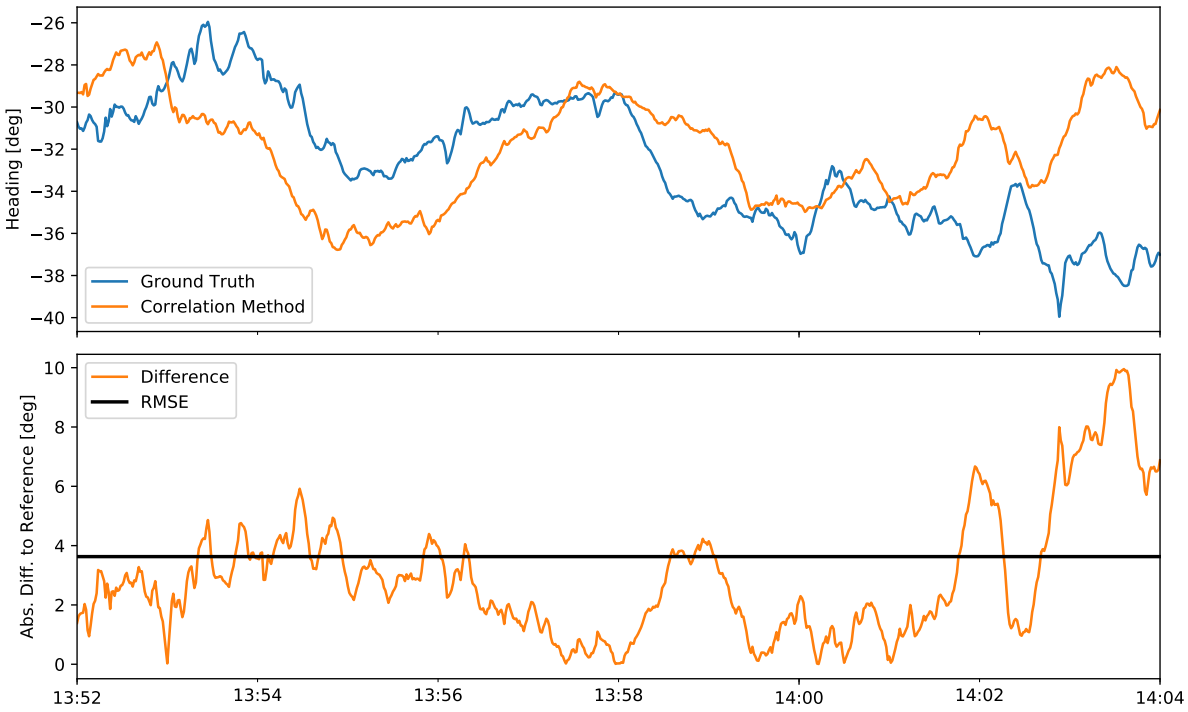



Figure 9.6: Comparison of results from correlation method to ground truth data in fourth lap

In comparison to the methods analyzed before, *stereo VO* is by far the most complex and computationally intensive of all. Figure 9.7 shows, that this method, despite the irregular shape of the windrow in the second lap, resembles ground truth data very closely. The RMSE over the course of a complete lap is only slightly higher than one degree, nearly meeting the additional goal defined before. Due to the nature of VO being a relative attitude determination algorithm, it is susceptible to drift. This

behaviour is perceptible over the whole timespan. It starts to become more evident from 12:34 onwards. Overall, the results of VO during the second lap fulfill nearly all requirements stated before. The only downside is that if one wants to use this technology in real-time to help a compost turner find the right path, a computation unit with enough power has to be present. As this thesis is just a feasibility study, the possibility for real-time implementation of the aforesaid algorithms is not a requirement.

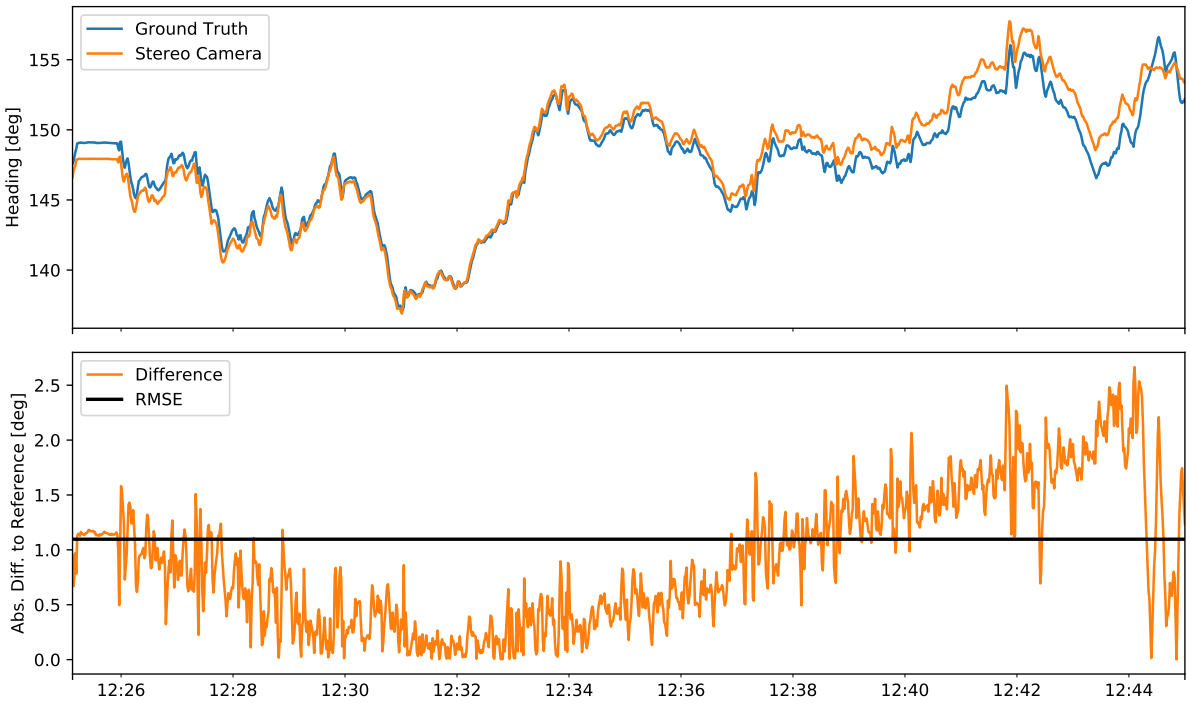



Figure 9.7: Comparison of results from VO to ground truth data in second lap

Data recorded during the fourth lap, resulted in the best representation of the ground truth data. The RMSE is well below one degree, fulfilling also the last requirement. Drift does not seem to be visible at first, because only the absolute differences to ground truth are visualized in Figure 9.8. When the upper part of Figure 9.8 is investigated more closely, one can see that data from the stereo camera at the beginning is more negative and during the end of the timespan the differences are positive. This concludes that drift is visible during all test runs, but with varying magnitude.

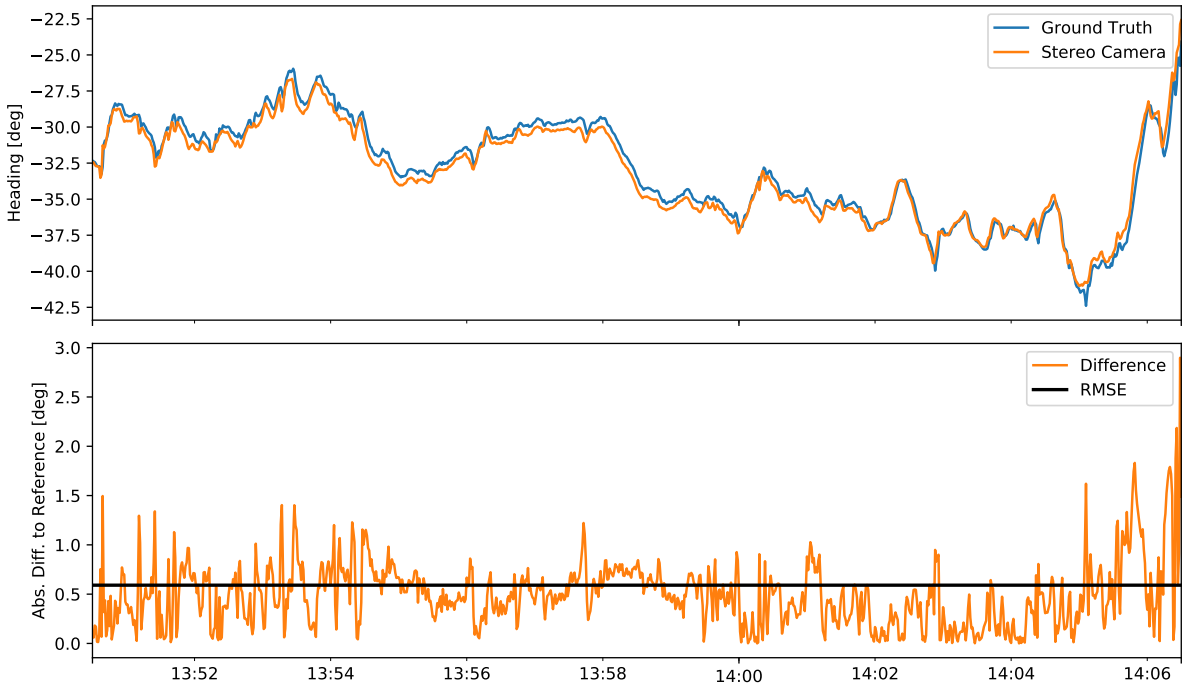



Figure 9.8: Comparison of results from VO to ground truth data in fourth lap

To summarize, the process of using image-based data in combination with the suited algorithms can yield a very robust, reliable and accurate method to gain attitude information of compost turners in challenging environments. As a passive technology, stereo vision systems are also immune to changing environmental conditions like temperature and humidity. The refractive properties of the media compost and it's surroundings did not influence the quality of the images and the derived depth information. A RMSE of less than one degree during a complete drive alongside a windrow has been achieved in the fourth lap. Therefore, VO fulfills all the requirements set. In conclusion, it can be said that by using a stereo camera as sensor, paired with a VO algorithm, precise relative heading determination of a compost turner can be accomplished. Table 9.2 gives an overview of all methods and their RMSE.

| **RMSE [deg]** | | |
|---|---|---|
| | **Lap 2** | **Lap 4** |
| Ultrasonic (simple approach) | 5.5* | 4.0* |
| Stereo Camera (correlation) | 8.8 | 3.6 |
| Stereo Camera (VO) | 1.1 | 0.6 |

Table 9.2: RMSE of all methods in comparison
*Value is only valid for a short timespan and not the complete lap

By using VO as an algorithm for pose estimation, one does not only get the attitude parameters *roll*, *pitch* and *yaw* and therefore the rotation matrix $\boldsymbol{R}$, but also a 3D translation vector $\boldsymbol{t}$. The complete pose $\boldsymbol{P}$, in matrix notation, can be given as

$$\mathbf{P} = \left\{ \begin{array}{c} \boldsymbol{R} \\ \boldsymbol{t} \end{array} \right\}, \qquad \text{where} \qquad \boldsymbol{R} = \boldsymbol{R}(r, p, y) \qquad \text{and} \qquad \boldsymbol{t} = \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix}.$$

To also compare the horizontal and vertical position components of the estimated pose, Figure 9.9 visualizes the horizontal components of the ground truth data (reference) and VO pose. The horizontal positioning errors after a complete lap are in the range of 1 - 1.5 m, which is very good for only using a relative method like VO. The accuracy could be improved by combining the relative measurements from the stereo camera with an absolute positioning technology like GNSS. The accuracy of the vertical component would greatly benefit from an accurate way of initializing the attitude of the camera, because it heavily depends on the pitch angle. This could be accomplished by using an IMU to initialize attitude.

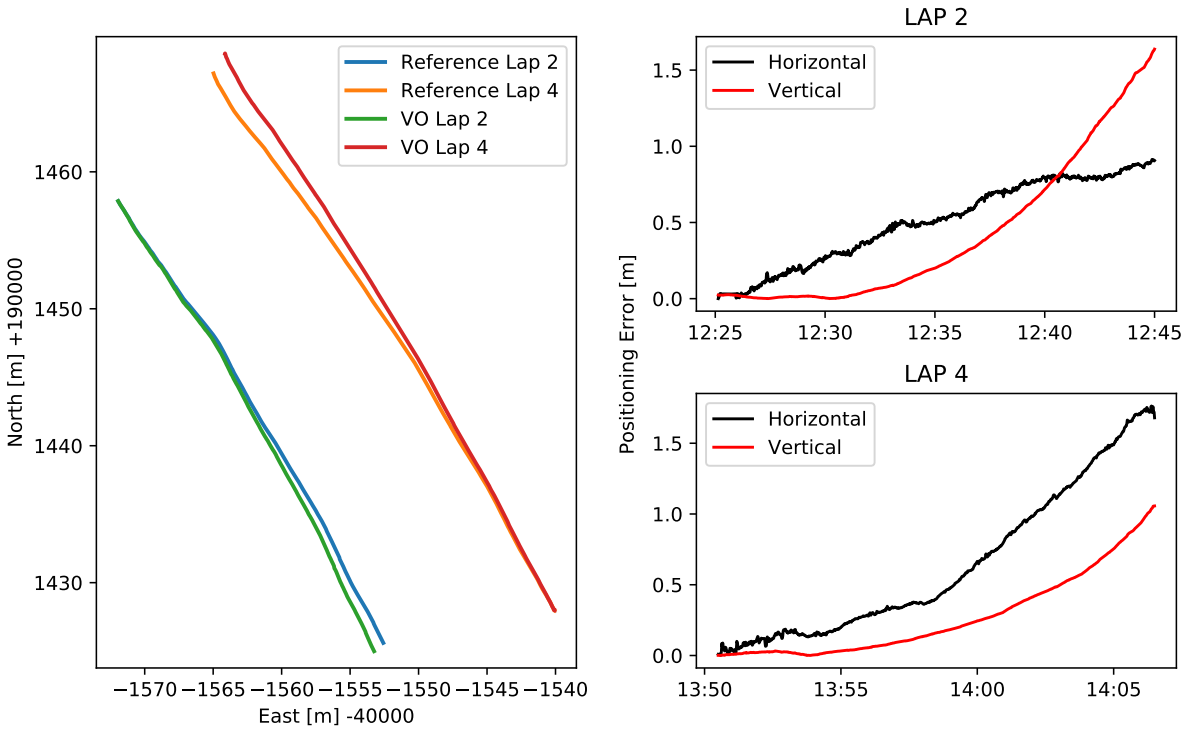



Figure 9.9: Left figure: Horizontal position plot of ground truth data and VO solution for laps 2 and 4; Right figures: Horizonal and vertical positioning error for laps 2 and 4

# 10 Conclusion

To guide a track-driven vehicle safely through a challenging environment, it is fundamental to know it's attitude parameters very precisely. This thesis tried to gather viable heading information from distance measurements taken on windrows below compost turners. The shape of a windrow stands out, because usually it is triangular. This property was taken and gives reason for an assumption, that it may be possible to use this information for heading determination in combination with the matching algorithm.

Three different sensor technologies have been selected and analyzed regarding accuracy, precision and robustness. The high accuracy and precision of laser distance measurement has been demolished by the poor resilience against dust formations in the air. According to literature, ultrasonic DME is a good choice for challenging and dusty environments, because of the wavelength used in the measurement principle. Unfortunately, this advantage of sonar technology could not be confirmed with the sensors used in this thesis. As a matter of fact, ultrasonic technology had the worst signal-to-noise ratio of all technologies used during the measurement campaigns. The stereo camera, on the other hand, was believed to have difficulties with dusty environments and particle formations in the air, hovering in front of the optics. This technology proved to be very reliable and accurate in generating depth maps, even when particle formations in the air were very dense.

Investigations regarding algorithms for heading determination from distance measurements ranged from simplistic to very sophisticated image-based approaches. It is believed, that the poor performance achieved, when creating accurate surface profiles with the help of spatial voting or binary morphism, was attributed to utilizing too little sensors in a non-optimal configuration. Extracting information about the location of the peak of a windrow with a depth-map correlation-approach seemed promising at first, but proved to be heavily dependent on the assumption of a perfectly triangle-shaped windrow. Surprisingly, it were the camera's images, paired with a sophisticated stereo VO algorithm, which yielded the best results of all methods tested. In the end,

it has been possible to achieve all the goals set at the beginning and in addition to that VO was able to achieve a very low RMSE throughout the test campaigns.

Although all requirements seem to be met, research on this topic is still at the very beginning. As mentioned before, a sensor array containing a larger number of individual sensors, with the possibility to change their arrangement angle on-the-fly, could lead to a more reliable and accurate way of creating surface profiles. One of the most interesting topics for the future is definitely the selection of a VO algorithm, perfectly suited for the tested environment. It would also be interesting to see how *feature-based* VO algorithms compare to *direct* ones. Feature detection methods could also be tailored to the homogeneous character of images taken from compost. Another advantage of image-based technologies is the utilization of Simultaneous Localization and Mapping (SLAM) algorithms. These combine the pose estimation of VO with a mapping algorithm, which creates a visual representation of the surroundings in real-time.

# Bibliography


[1] F. Zucconi, "Comparison of three windrow compost systems", 1982.

[2] NASA, *Scientific and Technical Aerospace Reports*. Scientific, Technical Information Office, National Aeronautics, and Space Administration, 1994, 768 pp., Google-Books-ID: gh5nttfWr5gC.

[3] M. Frank, "A step towards the design of collaborative autonomous machines : A study on construction and mining equipment", 2019. [Online]. Available: http://urn.kb.se/resolve?urn=urn:nbn:se:bth-18944 (visited on 01/29/2020).

[4] G. R. Opshaug and P. Enge, *Robotic snow cat*. 2000.

[5] S. Siebert. (Dec. 22, 2016). Bio-waste recycling in europe against the backdrop of the circular economy package, European Compost Network, [Online]. Available: https://www.compostnetwork.info/download/bio-waste-recycling-europe-backdrop-circular-economy-package/ (visited on 01/29/2020).

[6] M. Vogt, T. Neumann, and M. Gerding, "Frequency-diversity technique for reliable radar level measurement of bulk solids in silos", in *2013 European Radar Conference*, ISSN: null, Oct. 2013, pp. 129–132.

[7] Z. Zhang, S. Liu, G. Tsai, H. Hu, C.-C. Chu, and F. Zheng, "PIRVS: An advanced visual-inertial SLAM system with flexible sensor fusion and hardware co-design", in *2018 IEEE International Conference on Robotics and Automation (ICRA)*, ISSN: 2577-087X, May 2018, pp. 3826–3832. DOI: 10.1109/ICRA.2018.8460672.

[8] W. R. Blevin and W. J. Brown, "A precise measurement of the stefan-boltzmann constant", *Metrologia*, vol. 7, no. 1, pp. 15–29, Jan. 1971, ISSN: 0026-1394. DOI: 10.1088/0026-1394/7/1/003. (visited on 09/04/2019).

[9] O. Vaisala, "Humidity conversion formulas", 2013.

[10] M. Gilowski, C. Schubert, M. Zaiser, W. Herr, T. Wübbena, T. Wendrich, T. Müller, E. M. Rasel, and W. Ertmer, "Narrow bandwidth interference filter-stabilized diode laser systems for the manipulation of neutral atoms", *Optics Communications*, vol. 280, no. 2, pp. 443–447, Dec. 15, 2007, ISSN: 0030-4018. DOI: 10.1016/j.optcom.2007.08.043. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0030401807008577 (visited on 12/17/2019).

[11] J. Ryde and N. Hillier, "Performance of laser and radar ranging devices in adverse environmental conditions", *Journal of Field Robotics*, 2009.

[12] R. Szeliski, *Computer Vision: Algorithms and Applications*, ser. Texts in Computer Science. London: Springer-Verlag, 2011, ISBN: 978-1-84882-934-3. DOI: 10.1007/978-1-84882-935-0. [Online]. Available: https://www.springer.com/gp/book/9781848829343 (visited on 12/16/2019).

[13] F. A and B. C, "Random sample consensus", *Communications of the ACM*, Jun. 1, 1981. [Online]. Available: https://dl.acm.org/doi/abs/10.1145/358669.358692 (visited on 01/10/2020).

[14] B. Barshan, D. Baskent, and B. Barshan, "Comparison of two methods of surface profile extraction from multiple ultrasonic range measurements", *Measurement Science and Technology*, vol. 11, no. 6, pp. 833–844, Jun. 1, 2000, ISSN: 0957-0233, 1361-6501. DOI: 10.1088/0957-0233/11/6/330. [Online]. Available: http://stacks.iop.org/0957-0233/11/i=6/a=330?key=crossref.237958c4cf5e134dcb80830a7f90db36 (visited on 09/02/2019).

[15] E. R. DOUGHERTY, "An introduction to morphological image processing", *SPIE, 1992*, 1992. [Online]. Available: https://ci.nii.ac.jp/naid/10019616106/ (visited on 01/14/2020).

[16] P. Bedi, "Effective morphological extraction of true fingerprint minutiae based on the hit or miss transform", *International Journal of Biometric and Bioinformatics*, vol. 4, Jun. 1, 2010.

[17] D. Baskent and B. Barshan, "Surface profile determination from multiple sonar data using morphological processing", *The International Journal of Robotics Research*, vol. 18, no. 8, pp. 788–808, Aug. 1, 1999, ISSN: 0278-3649. DOI: 10.1177/02783649922066565. [Online]. Available: https://doi.org/10.1177/02783649922066565 (visited on 01/14/2020).

[18] F.-Y. Wang, "A simple and analytical procedure for calibrating extrinsic camera parameters", *IEEE Transactions on Robotics and Automation*, vol. 20, no. 1, pp. 121–124, Feb. 2004, ISSN: 2374-958X. DOI: 10.1109/TRA.2003.820919.

[19] K. Nagatani, A. Ikeda, G. Ishigami, K. Yoshida, and I. Nagai, "Development of a visual odometry system for a wheeled robot on loose soil using a telecentric camera", *Advanced Robotics*, vol. 24, no. 8, pp. 1149–1167, Jan. 1, 2010, ISSN: 0169-1864. DOI: 10.1163/016918610X501282. [Online]. Available: https://doi.org/10.1163/016918610X501282 (visited on 01/20/2020).

[20] A. Krause. (). Monocular visual odometry using a planar road model to solve scale ambiguity, The Robotics Institute Carnegie Mellon University, [Online]. Available: https://www.ri.cmu.edu/publications/monocular-visual-odometry-using-a-planar-road-model-to-solve-scale-ambiguity/ (visited on 01/20/2020).

[21] R. Mur-Artal and J. D. Tardos, "ORB-SLAM2: An open-source SLAM system for monocular, stereo and RGB-d cameras", *IEEE Transactions on Robotics*, vol. 33, no. 5, pp. 1255–1262, Oct. 2017, ISSN: 1552-3098, 1941-0468. DOI: 10.1109/TRO.2017.2705103. arXiv: 1610.06475. [Online]. Available: http://arxiv.org/abs/1610.06475 (visited on 11/21/2019).

[22] J. Engel, T. Schöps, and D. Cremers, "LSD-SLAM: Large-scale direct monocular SLAM", in *Computer Vision – ECCV 2014*, D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars, Eds., ser. Lecture Notes in Computer Science, Cham: Springer International Publishing, 2014, pp. 834–849, ISBN: 978-3-319-10605-2. DOI: 10.1007/978-3-319-10605-2_54.

[23] Y. Liu and C. Wang, "Hybrid real-time stereo visual odometry for unmanned aerial vehicles", *Optical Engineering*, vol. 57, no. 7, p. 073 104, Jul. 2018, ISSN: 0091-3286, 1560-2303. DOI: 10.1117/1.OE.57.7.073104. [Online]. Available: https://www.spiedigitallibrary.org/journals/Optical-Engineering/volume-57/issue-7/073104/Hybrid-real-time-stereo-visual-odometry-for-unmanned-aerial-vehicles/10.1117/1.OE.57.7.073104.short (visited on 01/20/2020).

[24] H. Wang, X. Wu, Z. Chen, and Y. He, "A novel hybrid visual odometry using an RGB-d camera", in *2018 33rd Youth Academic Annual Conference of Chinese Association of Automation (YAC)*, ISSN: null, May 2018, pp. 47–51. DOI: 10.1109/YAC.2018.8406345.

[25] S. Poddar, R. Kottath, and V. Karar, "Evolution of visual odometry techniques", *arXiv:1804.11142 [cs]*, Apr. 30, 2018. arXiv: 1804.11142. [Online]. Available: http://arxiv.org/abs/1804.11142 (visited on 01/20/2020).

[26] D. M. Patel and S. Upadhyay, "Optical flow measurement using lucas kanade method", 2013. DOI: 10.5120/9962-4611.

[27] J. Campbell, R. Sukthankar, and I. Nourbakhsh, "Techniques for evaluating optical flow for visual odometry in extreme terrain", in *2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (IEEE Cat. No.04CH37566)*, ISSN: null, vol. 4, Sep. 2004, 3704–3711 vol.4. DOI: 10.1109/IROS.2004.1389991.

[28] Y. Cheng, M. Maimone, and L. Matthies, "Visual odometry on the mars exploration rovers", in *2005 IEEE International Conference on Systems, Man and Cybernetics*, ISSN: 1062-922X, vol. 1, Oct. 2005, 903–910 Vol. 1. DOI: 10.1109/ICSMC.2005.1571261.

[29] K. Konolige, M. Agrawal, R. Bolles, C. Cowan, M. Fischler, and B. Gerkey, "Outdoor mapping and navigation using stereo vision", presented at the Proceedings of the International Symposium on Experimental Robotics, vol. 39, Jan. 1, 2006, pp. 179–190. DOI: 10.1007/978-3-540-77457-0_17.

[30] H. Schweinzer and P. Krammer, "Scene analysis with ultrasonic sensors", in *2005 IEEE Conference on Emerging Technologies and Factory Automation*, vol. 1, Catania, Italy: IEEE, 2005, pp. 219–226, ISBN: 978-0-7803-9401-8. DOI: 10.1109/ETFA.2005.1612523. [Online]. Available: http://ieeexplore.ieee.org/document/1612523/ (visited on 01/09/2020).

[31] J. H. Lambert and E. ( Anding, *Lamberts Photometrie : [Photometria, sive De mensura et gradibus luminus, colorum et umbrae] (1760)*, in collab. with Harvard University. Leipzig : W. Engelmann, 1892, 433 pp. [Online]. Available: http://archive.org/details/lambertsphotome00lambgoog.

[32] Hilti, *Hilti PD5 laser range meter*. [Online]. Available: www.hilti.com.

[33] A. Al Tahtawi, "Kalman filter algorithm design for HC-SR04 ultrasonic sensor data acquisition system", *IJITEE (International Journal of Information Technology and Electrical Engineering)*, vol. 2, Jul. 21, 2018. DOI: 10.22146/ijitee.36646.

[34] B. Hofmann-Wellenhof, K. Legat, and M. Wieser, *Navigation: Principles of Positioning and Guidance*. Wien: Springer-Verlag, 2003, ISBN: 978-3-211-00828-7. DOI: 10.1007/978-3-7091-6078-7. [Online]. Available: https://www.springer.com/gp/book/9783211008287 (visited on 02/17/2020).