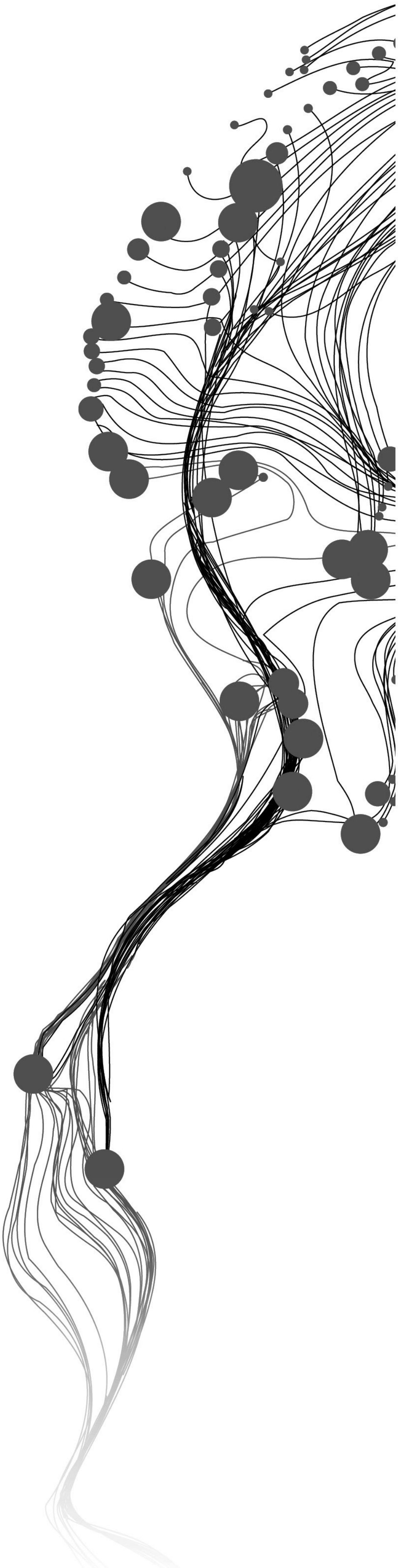
# WHEAT EAR DETECTION IN PLOTS BY SEGMENTING MOBILE LASER SCANNER DATA

KAAVIYA VELUMANI
February, 2017

SUPERVISORS:
Dr. ir. S. J. Oude Elberink
Dr. M. Y. Yang



# WHEAT EAR DETECTION IN PLOTS BY SEGMENTING MOBILE LASER SCANNER DATA

KAAVIYA VELUMANI
Enschede, The Netherlands, February, 2017

Thesis submitted to the Faculty of Geo-information Science and Earth Observation of the University of Twente in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation.
Specialization: Geoinformatics

SUPERVISORS:

Dr. ir. S. J. Oude Elberink
Dr. M. Y. Yang

THESIS ASSESSMENT BOARD:

Prof. Dr. ir. M. G. Vosselman (chair)
Dr. R. C. Lindenbergh; Delft University of Technology, Optical and Laser Remote Sensing

# ABSTRACT

The use of Light Detection and Ranging (LiDAR) to study agricultural crop traits is becoming popular. This is due to LiDAR's capability to render accurate 3-dimensional representation of the plant architecture. Wheat plant traits such as crop height, biomass fractions and plant population are of interest to agronomists and biologists for the assessment of a genotype's performance in the environment. Among these performance indicators, plant population in the field is still widely estimated through manual counting which is a tedious and labour intensive task. Thus, the goal of this study is to explore the suitability of LiDAR observations to automate the counting process by the individual detection of wheat ears in the agricultural field. However, this is a challenging task owing to the random cropping pattern and noisy returns present in the point cloud. The goal is achieved by first segmenting the 3D point cloud followed by the classification of segments into ears and non-ears. In this study, two segmentation techniques: a) voxel-based segmentation and b) mean shift segmentation were adapted to suit the segmentation of plant point clouds. A novel strategy was developed to distinguish the ear segments from leaves, stem and other plant organs. Finally, the ears extracted by the automatic methods were compared with reference ear segments prepared by manual segmentation.

The manual segmentation tests carried out with 6 operators revealed that it is hard even for humans to identify individual wheat ears from the point cloud. Also, the robustness of the two segmentation methods for detecting wheat ears over different crop developmental stages, wheat varieties and point densities was evaluated and compared. Both the methods had an average detection rate of 85%, aggregated over different flowering stages. The voxel-based approach performed well for late flowering stages (wheat crops aged 210 days or more) with a mean percentage accuracy of 94% and takes less than 20 seconds to process 50,000 points. Meanwhile, the mean shift approach showed comparatively better counting accuracy of 95% for early flowering stage (crops aged below 225 days) and takes approximately 4 minutes to process 50,000 points. Even though both the ear extraction approaches are dependent on point density, their performance was found to be consistent for up to 75% of the original point density of $16points/cm^2$. Thus, two ear detection methods, that use only the 3D coordinates of the plant canopy, have been developed. They can be extended to suit crops such as barley, millets, etc. that are structurally similar to wheat.



#### Keywords

*segmentation, wheat ear detection, voxelisation, mean shift, lidar*

# ACKNOWLEDGEMENTS

I would like to begin with a heartfelt thanks to my parents and brother for always encouraging me to pursue my passion.

I would like to thank my first supervisor, Dr Sander Oude Elberink, for always being supportive and available for discussions as the project progressed. His comments were crucial in shaping the thesis. I would also like to thank my second supervisor, Dr Michael Yang whose constructive criticism guided me to add extra components to my research work.

I take this opportunity to thank Dr Frederic Baret for his enthusiastic guidance and valuable inputs. Also my sincere thanks to Samuel Thomas and Benoit de Solan for providing the data necessary for the research and patiently answering my questions regarding the data acquisition.

I should also thank my dear friends who helped to prepare the reference datasets for this research work - manual segmentation of point cloud is not an easy task!! A special mention to Ranchana and Karma for spreading positive vibes when I most needed it.

Finally, I would like to thank all my lovely friends from GFM and Enschede for the amazing 18 months!

# TABLE OF CONTENTS


**Abstract** i

**Acknowledgements** ii

**1 Introduction** 1
- 1.1 Motivation and Problem Statement . . . 1
- 1.2 Research Identification . . . 2
  - 1.2.1 Research Objectives . . . 3
  - 1.2.2 Research Questions . . . 3
  - 1.2.3 Innovation Aimed At . . . 4
- 1.3 Project Set-Up . . . 4
  - 1.3.1 Project Work-flow . . . 5
  - 1.3.2 Thesis Structure . . . 5

**2 Literature Review** 7
- 2.1 Description of the Object of Interest . . . 7
- 2.2 Segmentation . . . 8
- 2.3 Individual Tree Detection . . . 8
- 2.4 Plant organ segmentation and Individual Plant detection . . . 9
- 2.5 Description of Short-Listed Methods . . . 10
  - 2.5.1 Voxel-based Connected Components . . . 10
  - 2.5.2 Mean-Shift Segmentation . . . 10

**3 Design of Methodology** 13
- 3.1 Data Preparation . . . 13
  - 3.1.1 Data Acquisition . . . 13
  - 3.1.2 Data Description . . . 14
  - 3.1.3 Pre-processing . . . 14
- 3.2 Segmentation . . . 15
  - 3.2.1 Voxel-Based Connected Components Segmentation . . . 15
  - 3.2.2 Mean-Shift Segmentation . . . 18
  - 3.2.3 Ear Classification . . . 20

**4 Parameter Optimization** 25
- 4.1 Sensitivity Analysis . . . 25
  - 4.1.1 Voxel-Based Ear Detection . . . 25
  - 4.1.2 Mean Shift-Based Ear Detection . . . 27
- 4.2 Parameter Tuning . . . 28
  - 4.2.1 Voxel-Based Ear Detection . . . 29
  - 4.2.2 Mean Shift-Based Ear Detection . . . 30
  - 4.2.3 Tuning over varying Point Density . . . 30

**5 Performance Evaluation** **31**
5.1 Evaluation Metrics . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 31
5.1.1 Root Mean Square Error . . . . . . . . . . . . . . . . . . . . . . 31
5.1.2 Mean Absolute Percentage Error . . . . . . . . . . . . . . . . . . 31
5.1.3 Qualitative metrics . . . . . . . . . . . . . . . . . . . . . . . . . 32
5.2 Evaluation Criteria . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 32
5.2.1 Effect of Developmental Stage . . . . . . . . . . . . . . . . . . . 33
5.2.2 Effect of Wheat Variety and Treatment . . . . . . . . . . . . . . . 35
5.2.3 Effect of Point Density . . . . . . . . . . . . . . . . . . . . . . . 36
5.3 Evaluation Summary . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 37

**6 Results and Discussion** **39**
6.1 Ear detection Results . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 39
6.1.1 Automatic Ear Detection Results . . . . . . . . . . . . . . . . . . 39
6.1.2 Manual Segmentation . . . . . . . . . . . . . . . . . . . . . . . . 40
6.1.3 Manual Segmentation Variability . . . . . . . . . . . . . . . . . . 41
6.1.4 Comparison with Manual Labels . . . . . . . . . . . . . . . . . . 41
6.2 Qualitative Observations . . . . . . . . . . . . . . . . . . . . . . . . . . . 43
6.2.1 Segmentation Quality . . . . . . . . . . . . . . . . . . . . . . . . 43
6.2.2 Processing Time . . . . . . . . . . . . . . . . . . . . . . . . . . . 44
6.3 Discussion . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 44
6.3.1 Usability of Intensity Information . . . . . . . . . . . . . . . . . . 44
6.3.2 Influence of Lidar Look Angle . . . . . . . . . . . . . . . . . . . . 45
6.4 Summary . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 46

**7 Conclusion and Recommendations** **49**
7.1 Conclusion . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 49
7.1.1 Research Questions: Answered . . . . . . . . . . . . . . . . . . . 49
7.2 Recommendations . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 51

**References** **53**

**A Number of Points per Ear** **57**
A.1 Method Description . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 57
A.2 Threshold Selection . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 57

# LIST OF FIGURES


1.1 The world population projection . . . . . . . . . . . . . . . . . . . . . . . . . . 1
1.2 Sample 3D point cloud of a wheat plot . . . . . . . . . . . . . . . . . . . . . . . . 3
1.3 Project Work-flow . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 6

2.1 Wheat spike at different flowering stages . . . . . . . . . . . . . . . . . . . . . . 7
2.2 Description of Skeletonisation . . . . . . . . . . . . . . . . . . . . . . . . . . . 10

3.1 Sensor set-up on the Mobile Laser Scanner . . . . . . . . . . . . . . . . . . . . . 13
3.2 Flow Diagram of Voxel-Based Segmentation . . . . . . . . . . . . . . . . . . . . 16
3.3 Flow Diagram of Mean Shift Segmentation . . . . . . . . . . . . . . . . . . . . 19
3.4 Flow Diagram - Ear Classification Strategy . . . . . . . . . . . . . . . . . . . . 21
3.5 Example to illustrate the steps involved in ear classification . . . . . . . . . . . . 22
3.6 Height histogram of the ear and non-ear segments . . . . . . . . . . . . . . . . 23

4.1 Sensitivity of Ear detection to the input parameters of Voxel-Based Approach . . 26
4.2 Sensitivity of the Ear detection to input parameters of the Mean Shift Approach 27
4.3 Experimental set-up for Parameter Tuning . . . . . . . . . . . . . . . . . . . . . 28

5.1 Effect of crop developmental stage on the automatic ear detection methods . . . 33
5.2 Effect of Wheat Variety on Ear Detection . . . . . . . . . . . . . . . . . . . . . 35
5.3 Effect of point density on ear detection . . . . . . . . . . . . . . . . . . . . . . 36

6.1 Voxel-Based Ear Detection results . . . . . . . . . . . . . . . . . . . . . . . . . 39
6.2 Mean Shift-Based Ear Detection results . . . . . . . . . . . . . . . . . . . . . . 40
6.3 Ear Counts from Manual and Automatic Detection Methods . . . . . . . . . . . 42
6.4 Comparison of Quality of Segments . . . . . . . . . . . . . . . . . . . . . . . . 43
6.5 Corrected Intensity Histogram . . . . . . . . . . . . . . . . . . . . . . . . . . . 45
6.6 Data acquired from Nadir vs Lateral Look Angle . . . . . . . . . . . . . . . . . 46

A.1 Threshold selection for filtering the point cloud . . . . . . . . . . . . . . . . . . 58

# LIST OF TABLES

3.1 Description of the wheat plots used in the study . . . . . . . . . . . . . . . . . . 14
3.2 Parameters used in the voxel-based Segmentation . . . . . . . . . . . . . . . . . . 18
3.3 Parameters used in the Mean Shift Segmentation . . . . . . . . . . . . . . . . . . 20
3.4 Parameters used in classification of the segments . . . . . . . . . . . . . . . . . . 23

4.1 Optimal Parameters for voxel-based ear detection . . . . . . . . . . . . . . . . . . 29
4.2 Optimal parameters for mean shift based ear detection . . . . . . . . . . . . . . . 30
4.3 Optimal Parameters for different Point Densities . . . . . . . . . . . . . . . . . . 30

5.1 Effect of crop developmental stage on ear detection . . . . . . . . . . . . . . . . 34
5.2 Effect of Wheat Variety on Ear Detection . . . . . . . . . . . . . . . . . . . . . . 36
5.3 Effect of Point Density on Ear Detection . . . . . . . . . . . . . . . . . . . . . . 37

6.1 Count of Ears extracted by manual segmentation . . . . . . . . . . . . . . . . . . 41
6.2 Variability among operators observed during manual segmentation . . . . . . . 41
6.3 Error in Ear Counts from the Automatic Detection Methods . . . . . . . . . . . 42

# Chapter 1

# Introduction

## 1.1 MOTIVATION AND PROBLEM STATEMENT

Wheat, one of the major staple crops utilised mainly for food and feed, contributes to about 29% of the world's cereal produce (FAO, March 2016). In order to meet the demands of the projected world population of 9.6 billion in 2050 as seen in Figure 1.1, wheat production will have to be doubled (Scott, 2014). This may be achieved by the identification and cultivation of high yielding wheat varieties that can adapt to the climate change and perform well under different stress conditions. Hence, high-throughput plant phenotyping techniques (to study the observable traits of plant varieties in controlled or field environment) are needed to assess the performance of the crop varieties under varied treatments of irrigation and fertilizers (Furbank & Tester, 2011).

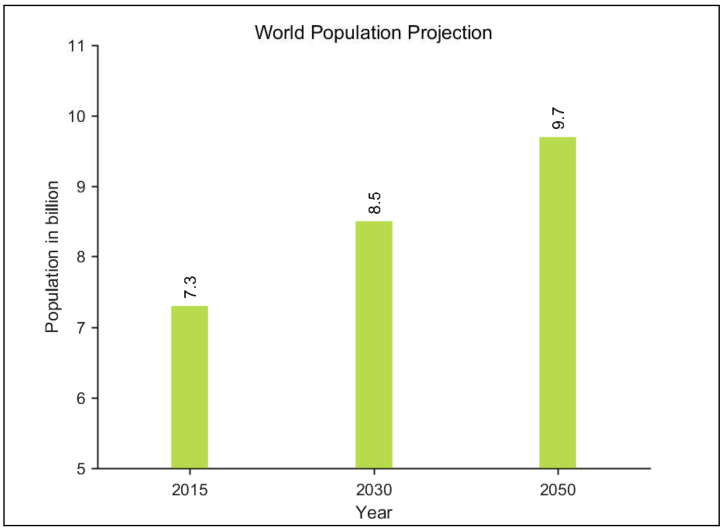



Figure 1.1: The world population projection as forecasted by the United Nation (UN DESA, 2015)

LiDAR (Light Detection And Ranging) or laser scanning technology has been identified to hold high potential for meeting the demands of next generation phenotyping challenges (Lin, 2015). This is attributed to the availability of LiDAR systems with small footprint and high pulse emission frequency and their capability to provide high-throughput plant traits. These systems provide robust data in varied illumination conditions and effective reconstruction of the in-field 3D crop architecture. Understandably, the use of laser scanners for field crop monitoring, both stationary and robot-mounted scanners, is becoming common for crop-height measurement and biomass estimation (Koenig et al., 2015; Hofle, 2014; Garrido et al., 2015). However, these applications have not yet exploited the full potential of LiDAR data i.e. 3D point clouds. Further, it could be extended to yield prediction.

The yield of a wheat plot can be determined by the number of grains per ear (i.e. spike) and the number of ears per $m^2$ which is highly correlated with the carbon and nitrogen availability in the crop (Sinclair & Jamieson, 2006). The number of ears per plot can be obtained by manual, semi-automated and automated techniques. Even though it is favourable to incorporate the automatic counting of wheat ears, it is a challenging task due to the random cropping pattern with close plant spacing and high extent of overlap (LemnaTec, 2015).

Few automated image processing techniques have been proposed to count wheat ears using 2D images from Charge-Coupled Device (CCD) cameras by applying texture-based classification in hybrid space (Cointault et al., 2008) or by high pass Fourier filtering (Journaux et al., 2010). The counting accuracy from 2D cameras is constrained by the illumination conditions at the time of image acquisition and undetected wheat ears due to overlap and obstruction by other plant organs. Deery et al. (2014) showcase the use of rasterised elevation images to count the number of wheat ears by applying a simple particle count algorithm on the segmented image. This approach does not perform well for fields with high crop density and overlapping spikes. Thus, in several places, manual counting is widely in practice. However, it is a tedious and labour intensive method and obviously cannot meet the needs of high-throughput phenotyping. Hence, 3D point clouds from LiDAR sensors could be used as an alternative for the automatic counting of wheat ears as they may help to overcome the limitations of the existing methods owing to the availability of the depth information and reliability in all illumination and atmospheric conditions.

Although several studies have been conducted on the derivation of field-crop parameters from LiDAR observations (Lin, 2015; see also Hosoi & Omasa, 2009, Houldcroft et al., 2005; Eitel et al., 2014), comparatively very few exist on the detection of individual crops in the field (Hofle, 2014; Weiss & Biber, 2011). Moreover, the extraction of plant organs from LiDAR observations in the field still remains a challenge. Thus, this study aims to explore the suitability of laser scanned data, acquired from sensors fitted in an automated vehicle, for estimating the number of wheat ears in an agricultural plot.

## 1.2 RESEARCH IDENTIFICATION

The use of robotics assisted imaging platforms have become popular in agricultural research in the past few years. Parallelly, interest towards the use of 3D data to study field crops has also been on the rise. This research explores the suitability of data from laser scanner, mounted on a mobile platform, to estimate wheat crop density by counting the number of ears.

Figure 1.2 shows a sample LiDAR observation acquired over a wheat plot, close to the harvest stage, colour coded with respect to height. From this figure, it is observed that it is hard even for a human eye to distinctly separate one ear from the other. Further, the different size and orientation of wheat ears within the same plot is observed. The fact that the geometry and size of wheat ears change with the age of the crop makes the development of an algorithm that is ideal for all developmental stages a challenge.

Thus, this research addresses the development of a suitable method - segmentation followed by classification - that uses the approximate geometry of the wheat ear to automatically extract them from the 3D point clouds. In addition, experiments were done to test the suitability of the intensity information recorded by the laser scanner to improve the segmentation process. The influence of point density on the segmentation accuracy is also evaluated. Finally, an evaluation of the performance of the designed method on different developmental stages of the plant is also presented. Developing a successful working method will be useful to agronomists and plant biologists in wheat yield prediction and individual crop monitoring. The algorithm can also be implemented on-line on combine harvester systems to improve their performance by controlling their speed depending on the plant density estimated.

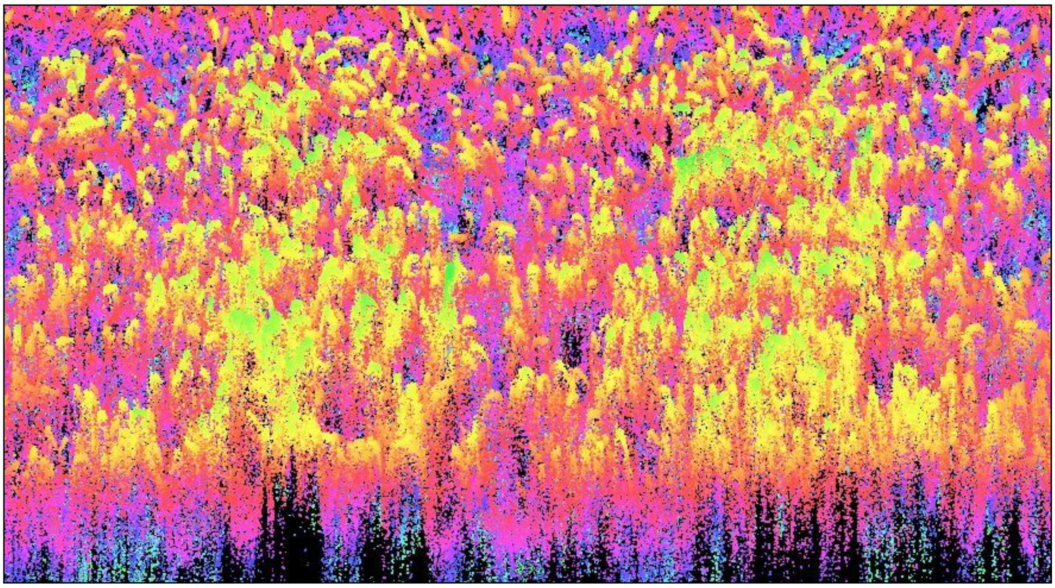


Figure 1.2: A sample 3D point cloud acquired over a wheat plot. The random spacing between plants, irregular orientation of the ears and noisy air returns make individual ear detection a challenge.

### 1.2.1 Research Objectives

The overall objective of the proposed study is as follows:

- To explore the possibility of counting the number of wheat ears in an agricultural field by processing laser scanned point clouds.

In-order to meet the overall objective, the following sub-objectives must be satisfied.

1. To perform pre-processing and reduce the noise in the data.
2. To design an appropriate segmentation method that can detect the wheat ears and the subsequent automation of the process.
3. To evaluate the influence of acquisition angle (lateral or nadir), suitability of intensity information from the laser scanner, wheat crop variety and the developmental stage on the segmentation process.

### 1.2.2 Research Questions

The following research questions should be answered
**Sub-objective 1:**

- What are the methods that have already been proposed in literature for the segmentation of plant organs?
- What are the methods in literature for single tree detection from airborne laser scanned data that are relevant to the proposed study?
- How can the knowledge on how wheat seeds are sown (the number of rows, average spacing between each plant and rows) and how they look like be incorporated in the segmentation algorithm?

**Sub-objective 2:**

- What is the size and distribution of the noise and how can it be reduced?
- What is the influence of the shadow returns (resulting in noisy objects) on the segmentation process?

**Sub-objective 3:**

- How does the performance of the segmentation algorithm differ depending on the angle of acquisition (nadir and lateral)? Which dataset (nadir or lateral) is preferable to estimate the count of ears?
- How does the point density affect the segmentation performance?
- Will the addition of intensity information from the laser scanner complement the counting accuracy?
- What is the effect of the variety and senescence of wheat ears on the segmentation performance?

### 1.2.3 Innovation Aimed At

The proposed study aims to develop a pipeline for the counting of wheat ears by automatic classification of point cloud acquired in an agricultural field. Previous studies have identified methods that can automatically segment plant organs from 3D point cloud of single plant acquired in controlled conditions i.e. at the laboratory and the greenhouse-scale. The innovation of this study is that it develops and compares two reliable approaches that detects and segments the wheat plants at field-scale; hence helping to overcome the hurdles faced with manual counting for the evaluation of genotype performance in the field. A novel classification strategy for distinguishing ear segments is proposed. The ear detection methods proposed are adaptable to other plants like barley, millets, etc. which are structurally similar to wheat crop. This study makes investigations to provide a detailed report on the robustness of the proposed methods for different experimental settings and how the methods could be adapted to different scenarios of point density and ear developmental stage.

## 1.3 PROJECT SET-UP

The research project will be carried out in four phases:

(A) Review and evaluation of existing methods

(B) Design and implementation

(C) Parameter optimisation

(D) Performance evaluation and comparison

### 1.3.1 Project Work-flow

In the first phase, a review of the point cloud segmentation techniques reported in literature was carried out. Among the numerous segmentation methods available, special attention was given to the ones that have been demonstrated on non-homogeneous vegetation structures. This aided in narrowing down segmentation methods that would be capable of clustering wheat ears in dense point clouds acquired over crops that lack a regular structure or sowing pattern. From the literature review, voxel-based segmentation and mean shift segmentation appeared the most promising to segment plant ears owing to their flexibility over different applications.

The second phase was devoted to the development of a methodology to estimate the wheat ear population. This was split into two steps:

(i) Segmentation of point clouds: The two segmentation methods short-listed in the previous phase were modified and adapted to eliminate outliers, multi-edge effect and identify our object of interest i.e. wheat ears.

(ii) Classification of segments: A new grammar was developed to classify the resulting plant organ segments from the first step into ears and non-ears.

In the third phase, we performed an initial sensitivity analysis of the input parameters used in the segmentation and ear classification procedure. This was followed by a cross validation technique to optimise the values for the input parameters for varying point densities by downsampling the datasets used in the study.

The final step was to evaluate the performance of the designed method on three different point densities, developmental stages and varieties of wheat. This gives a measure of the robustness of the designed strategy and demonstrates if it is sensitive to a certain dataset, point density, variety or developmental stage of the crop. Figure 1.3 describes the general methodology followed during the project execution.

### 1.3.2 Thesis Structure

The thesis content is organized in seven chapters. The first chapter gives a brief introduction on the motivation behind the study, the problem statement and the questions that will be addressed in the research. The second chapter deals with a brief review of the relevant segmentation methods found in literature. The third chapter describes the design of the methodology work-flow and implementation of two methods that could segment wheat ears in the canopy. The fourth chapter is on the optimisation of the input parameters used in the ear detection methods designed in this study. It also includes sensitivity analysis and devising strategies for the automatic selection of thresholds for different point densities. The fifth chapter is on the evaluation and comparison of the robustness of the proposed methods over different developmental stages, point density and varieties. The sixth chapter presents the results of ear detection and validates with reference ear segments extracted by manual segmentation. The final chapter contains a brief recap of the results, conclusion and the sections that could be addressed in future work.

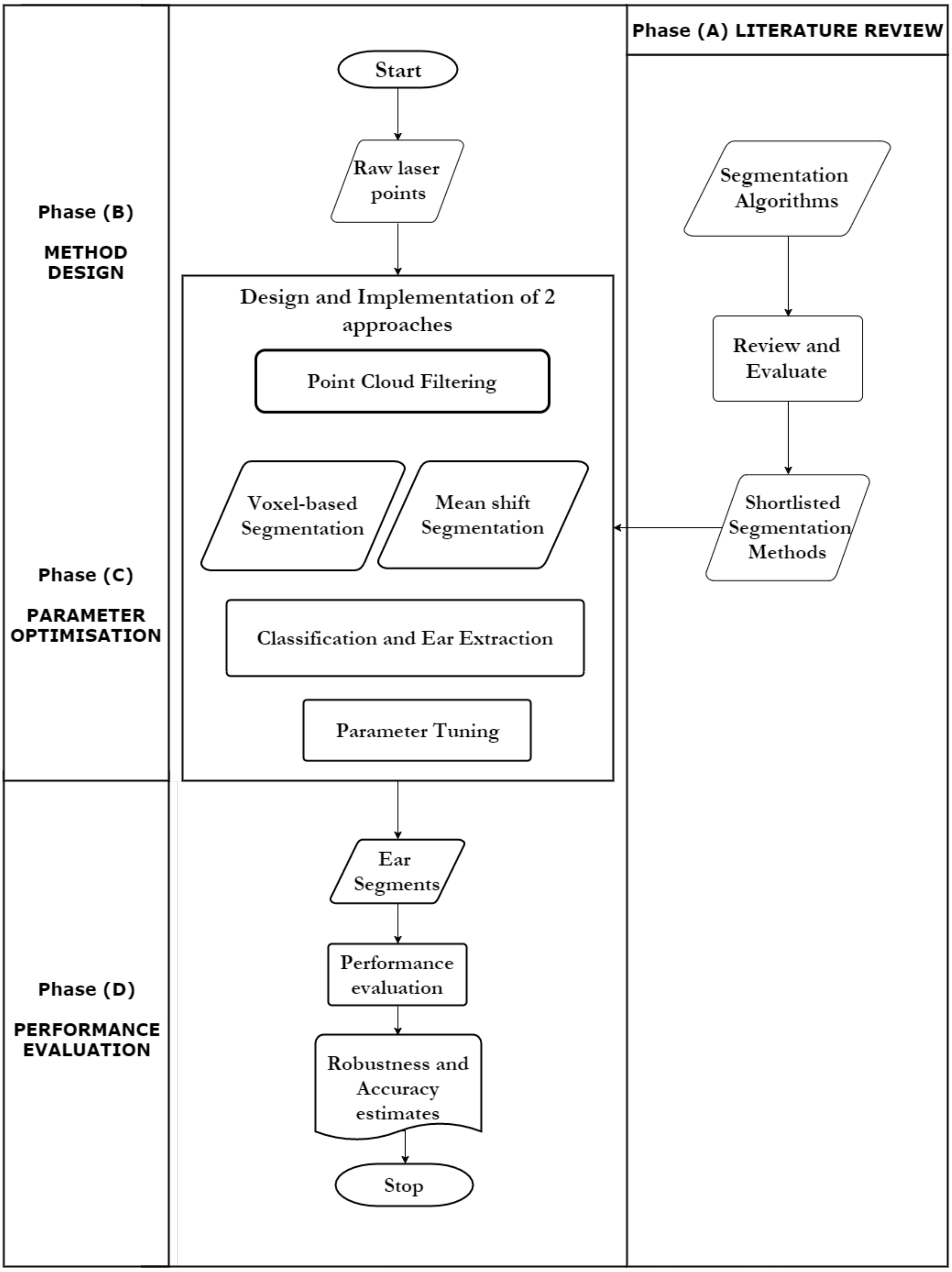



Figure 1.3: Flowchart depicting the project work-flow

# Chapter 2
# Literature Review

This chapter is a brief review of the existing segmentation methods that are relevant to the problem considered in this research. The section 2.1 describes the object of our interest i.e. the structure of wheat ears. The following section 2.2 focuses on the categories of the segmentation algorithms available, followed by a review of the segmentation approaches proposed in literature for detection of individual trees from laser scanned data. There is also a brief review on the methods available for the automatic classification of plant organs and extraction of plant parameters from 3D data. The final section 2.5 gives a description of the chosen methods that will be applicable for the problem statement with appropriate reasoning.

## 2.1 DESCRIPTION OF THE OBJECT OF INTEREST

For the problem under consideration, wheat spikes need to be distinguished from a cloud of stems, leaves and ground. Each stem of the wheat plant has a single protruding spike. Hence, the wheat spikes are found in the top layer of the canopy. The surface of the wheat spike is non-planar and is usually thin and long, approximating the shape of a prolate. Wheat spikes are green and erect in the early flowering stage and become brown and bent during the harvest period as is shown in Figure 2.1(a) and (b) (Miller, 1992). It should be noted that wheat is usually planted with irregularly spacing as can be seen in Figure 2.1(c). Depending on the variety of wheat plants and the developmental stage, the extent of overlap among adjacent plants varies. While designing an extraction method, these characteristics of the wheat canopy must be taken into consideration.

(a) Early flowering stage; green & erect spike (Miller, 1992).

(b) Late flowering stage; brown and bent spike (Knapton, 2016).

(c) A plot of wheat, where the irregular spacing between plants is noticeable (Knapton, 2016).

Figure 2.1: Wheat spike at different flowering stages showing different size and orientation of the ears.

## 2.2 SEGMENTATION

Point cloud segmentation may be defined as the clustering of points based on their properties and spatial distribution to form homogenous regions. In most scenarios, segmentation is an integral step to recognize objects in the point cloud which determines the amount of useful information retrieved (J. Wang & Shan, 2009). Hence, the application and the objects present in the point cloud should determine the choice of segmentation technique.

There are two basic design mechanism to segment point clouds. The first approach is based on methods that have mathematical model assumptions or geometric reasoning such as model fitting, probability density estimators or region growing. Though these methods achieve quick results for simple scenarios, they are sensitive to noise and exhibit poor performance in complex scenarios. The other approach is based on machine learning algorithms trained to classify the different object types in the scene (Nguyen & Le, 2013). A review of various segmentation techniques has been presented by Vosselman et al. (2004) and they are categorized based on the surface being extracted. For the extraction of smooth surfaces, they have suggested region growing, scan line segmentation and connected components in voxel space.

## 2.3 INDIVIDUAL TREE DETECTION

The detection of individual trees from LiDAR observations of forest canopy is popular for forest inventorying and monitoring. This task of detecting individual trees from airborne laser scanned data is comparable to the detection of individual ears from the wheat canopy. This is because, in both applications, the detected objects are non-homogeneous structures observed from an airborne perspective. The scale of observation and size of the detected objects are two criterion that differ. Hence, reviewing the commonly used techniques for tree detection would help to design the methodology to extract ears. There is extensive literature available for the detection of individual trees from LiDAR observations. Among them, the ones most suitable for the problem at hand are discussed in the following paragraphs.

Vauhkonen et al. (2012) compared the performance of a selection of segmentation techniques on different tree types. According to their study, detection of trees using local maxima with residual height adjustment (Solberg, Naesset, & Bollandsas, 2006) and geometric crown model based segmentation (Vauhkonen2012) resulted in good detection rates. It was concluded that the detection rates of the techniques varied depending on the tree type, tree density and clustering.

Another new segmentation technique that uses the relative spacing between trees at different height levels of the canopy to determine individual tree segments (Li et al., 2012). This technique is a top-to-bottom region growing approach that iteratively assigns a point to a tree segment after considering the distance between each point and its closest neighbouring tree point. The authors had also proposed a way to use the spatial distribution of points projected onto a 2D convex hull to determine if they belong to an extended branch or a neighbouring tree.

Sirmacek and Lindenbergh (2015) proposed a fast method that utilises a grid-based point density to assign probability values to the grid surface which are subsequently used in tree detection. In another study,Yao, Krzystek, and Heurich (2012) combined stem detection with normalized cut segmentation, a technique that has its roots in image processing in order to detect individual trees from laser scanner data. The method effectively detects trees, including the ones in the understory of the canopy. However, the drawback of applying these methods for the detection of wheat ears is that they employ the use of a local maxima as approximate tree positions. This would be misleading for wheat plants due to bent ears and the noisy multipath returns from neighbouring plant structures.

Mean-shift approach is another method widely used for detecting individual trees in urban environment (Weinmann, Mallet, & Brédif, 2016; Melzer, 2007) and for segmentation of different levels of forest structure (Ferraz et al., 2010).

A voxel approach followed by connected components for the detection of trees from mobile laser scanner data was presented by Gorte, Oude Elberink, Sirmacek, and Wang (2015) as one of the submissions to an IQmulus contest. Another voxel approach was developed by Y. Wang, Weinacker, and Koch (2008). It detects trees from canopy and sub-canopy layer by first normalizing the height of LiDAR observations using a DEM, then resampling them into a voxel space and finally searching for tree crowns after rasterizing the points in the voxel at different height levels.

## 2.4 PLANT ORGAN SEGMENTATION AND INDIVIDUAL PLANT DETECTION

Several authors have proposed different strategies for the automated reconstruction and segmentation of plant organs from close-up laser scanned point clouds. Paulus et al. (2013) used a surface feature histogram based approach for the automatic segmentation of plant organs that gave reliable results for different plants including grapevine and wheat. Another study by Paulus et al. (2014) used surface feature histogram accompanied with description of organ morphology for the automatic parameterization of 3D point clouds of barley organs. Both these studies employ a segmentation approach that uses region growing combined with Support Vector Machine (SVM) classifier. Wahabzada et al. (2015) demonstrated the use of an unsupervised clustering algorithm to achieve automatic segmentation of plant organs using unlabelled data. Another top-down partitioning approach that segments each plant organ in different steps was proposed by Paproki et al. (2011). All these methods have been demonstrated for close-up scan of individual plants in laboratory conditions. But the field conditions are very different in terms of crop density with overlapping plant structures and dusty environment. Hence, these methods cannot be directly implemented on laser data scanned in the field.

In comparison, very few works have been carried out in the detection of plants in the field. Weiss and Biber (2011) developed a strategy for real-time discrimination of plant from ground and the mapping of plant clusters in the field to aid in navigation of farm robots. In another study, Hofle (2014) developed a workflow for the detection of individual maize plants from laser scanned point clouds by applying an object based classification that uses geometric and radiometric features. The detection rates were high since the maize plants in this study were in early developmental stage (10 to 20 cm high) with minimal overlap and were planted at a predefined distance from each other. In another study, use of terrestrial laser scanner to estimate wheat crop density was demonstrated by Lumme et al. (2008). A detailed methodology was not presented and the authors conclude that a laser scanner mounted on a mobile platform could be used as a tool in precision farming. Saeys et al. (2009) developed a method to estimate wheat crop density by 3D reconstruction of an artificial canopy set-up. They extracted the ear layer by fitting a "thin-plate smoothing spline" and generated a point density image from which the approximate location of the ears was identified. This method was found to be computationally intensive due to the fitting of the spline. Also, the method was demonstrated on an artificial canopy, where overlap among the adjacent plants is minimal. Hence, it is still a challenge to make wheat ear detection from fully grown canopy operational at field scale. Also, a detailed comparison of wheat ear detection method that segment ears from 3D point clouds acquired over grown wheat plants (80 to 90 cm high),with high plant density resulting in high degree of overlap is not reported in literature.

## 2.5 DESCRIPTION OF SHORT-LISTED METHODS

Based on the literatures reviewed in the above sections, priority was given to methods that are flexible, do not require labelled training samples and are capable of identifying irregular structures from dense point clouds. The following two methods were thus chosen to make an initial experimentation.

- Voxel-based Connected Components
- Mean-Shift Segmentation

The following subsections present in detail, the existing works that lead to the selection of these two methods.

### 2.5.1 Voxel-based Connected Components

Voxel-based tree detection (Hosoi & Omasa, 2006) has been demonstrated for the estimation of forest parameters. Another popular application of voxelization of 3D point cloud is to extract the skeleton of trees (Bucksch et al., 2009) and plants (Ramamurthy et al., 2015) as in Figure 2.2. These works demonstrate how the voxel approach can be adapted to fit the object of interest.

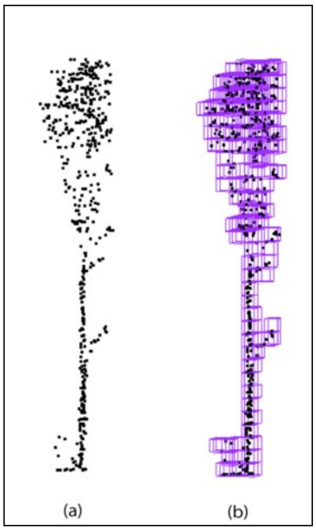



Figure 2.2: (a) Delineated tree points (b) Delineated points displayed with the voxel-cubes organised in octree organizations (Bucksch et al., 2009)

In this approach, the point cloud is split into equal sized voxel-cubes and the number of points within each voxel-cube is calculated. Based on the point density and the objects being studied, only the voxels with number of points above a threshold are considered for further processing. This approach could be used to delineate individual plants in the canopy layer by removing overlapping points from adjacent plants and segment the ears by using a connected component analysis.

### 2.5.2 Mean-Shift Segmentation

The mean shift approach (Fukunaga & Hostetler, 1975) uses a non-parametric density estimator function to look for modes in the data. This was first used in computer vision by Comaniciu and Meer (2002) where they demonstrated the applicability of mean-shift in image segmentation, to look for clusters in feature space of images. Since then, this method has gained popularity and has also been demonstrated on 3D point clouds.

Mean shift when applied to Lidar data looks for clusters in the 3D space. For each point in the dataset, it estimates a weighted mean of points that fall inside a kernel window, and shifts the centre of the kernel to this newly estimated mean. The procedure is repeated until convergence of the mean occurs and ensures quick convergence in areas with low point density. Since it is a non-parametric technique, it does not assume that the data fits a pre-defined distribution. The only user-defined parameter is the kernel bandwidth that could be defined based on the characteristics of the object being segmented (Melzer, 2007). Two main advantages of using mean-shift over k-means clustering is that

- there is no need to have a priori knowledge on number of clusters in the dataset
- it does not restrict itself to a spherical search space; the clusters can take any shape depending on the kernel bandwidth specified.

Thus, the mean-shift technique carried out using a Gaussian kernel (to ensure that the weights to the points are assigned in a Gaussian manner) with different bandwidths for X, Y and Z coordinates specified such that $X=Y<Z$, to approximate the shape of wheat ear would be successful in extracting the ear segments.

# Chapter 3
# Design of Methodology

The procedure to extract the wheat spikes/ears from the point cloud involves two main steps as follows:

(A) Segmentation of the point cloud

(B) Classification of the segments

One of the two short-listed segmentation techniques was applied on the point cloud. A common methodology was developed to classify the resulting segments into ears and non- ears. Finally, the wheat ear count from the two segmentation techniques was analysed and compared. The following sections in this chapter describe the steps for data pre-processing, segmentation and classification procedures that were carried out.

## 3.1 DATA PREPARATION

### 3.1.1 Data Acquisition

The datasets for this study were acquired from sensors mounted on an unmanned ground vehicle (UGV). The UGV was fitted with three SICK LMS 400 LiDAR (SICK Germany, 2007) sensors and two RGB cameras. Figure 3.1 depicts how the sensors were set-up in the acquisition platform. LiDAR 1 was mounted with an inclination of 45° whereas LiDAR 2 and LiDAR 3 were mounted

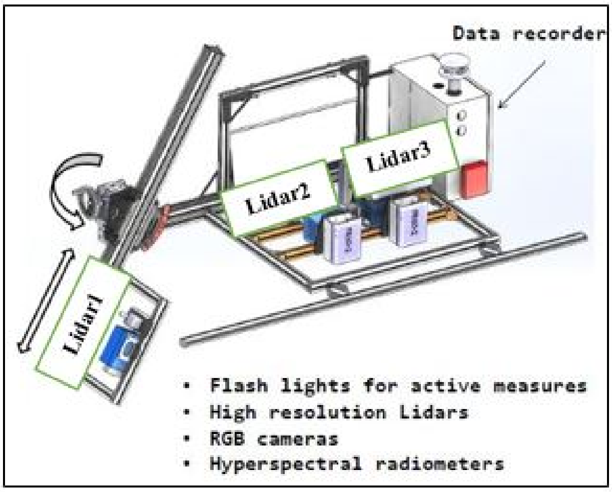



Figure 3.1: Description of the sensor positions on the mobile laser scanner (Description and image as provided by INRA, Avignon)

in the nadir, looking downward, separated by a distance of around 45 cm. This time of flight laser scanner has a scanning frequency of upto 500 Hz with an operating range of 0.7 m to 3 m and systematic error of $\pm 4$ mm. The scanner uses visible red light at 650 nm as its light source. The Camera 1 and Camera 2 were fitted next to LiDAR 1 and LiDAR 2 respectively. The datasets used in this study were acquired and provided by the French National Institute for Agronomic Research (INRA), Avignon.

### 3.1.2 Data Description

The laser scanning survey was conducted on three micro-plots, roughly of size 10 m x 2 m, in Gréoux, France. Three plots sown with three different varieties of wheat on 29 October 2015 and subject to different irrigation treatments were selected for the survey. In-order to have a time series, the survey was conducted on three different dates with the crops aged 194 days, 209 days and 225 days (May 10; May 25; June 10 2016 respectively), each 15 days apart, on the same micro-plots. Table 3.1 displays information regarding the wheat variety and irrigation treatment of the plots used in this study.

Table 3.1: Description of the wheat plots used in the study. (Data description as provided by INRA, Avignon).

| Plot | Plot ID | Cultivar | Irrigation | Number of Plants/$m^2$ |
|---|---|---|---|---|
| A | IVeuro 216 | SCULPTUR | Yes | 320 |
| B | IVeuro 406 | SERI*3//RL6010 | No | 277.1 |
| C | IVeuro 514 | TE1202 | No | 361.4 |

Along with the LiDAR observations, wheat spike density estimates from manual counting in the field were also provided. These estimates were used as reference to evaluate the counting results from the developed method. To check the segmentation accuracy, reference wheat spike segments were extracted and labelled manually using CloudCompare.

The data from the laser scanner is in the form of a dense point cloud stored as a long list of (X, Y, Z, I) points that gives the 3D position and recorded intensity value for the objects observed. The points were recorded in a local coordinate system initialized at (X=0, Y=0, Z=Z) for the first position acquired. The subsequent points were assigned co-ordinates based on (X positive right, Y positive forward and Z positive upward). The datasets provided were acquired from a height of $2.1m$ above the ground and had a point density of $16pts/cm^2$.

### 3.1.3 Pre-processing

On visual analysis, the point cloud was found to contain noisy points hovering above the canopy and below the ground level. Also, there were points floating around the objects in the scene that could be identified as "air returns" or "ghost returns" owing to the mixed edge effect. As per Van Genchten et al. (2008), the mixed edge effect occurs when a laser beam is intercepted by the edge of an object. The beam splits and thus two signals reflected by two different objects are sent back. The receiver records an averaged distance between the two received signals and stores it as a point, which does not actually exist. In our case-study, the presence of "ghost returns" was due to the mixed-edge effect owing to the beam divergence varying with range and the densely occurring leaves, spikes and soil particles.

To remove the above mentioned unwanted points, an initial cleaning was carried out. The point cloud was subject to a simple cleaning process where the negative height points and sensor locations were removed. The "ghost return" points surrounding the objects and hovering over the crop canopy were removed in the first step of the segmentation process. This removal was carried out based on the neighbouring point density, and was handled in different ways for the two short-listed segmentation techniques. The length of a wheat spike in the harvest stage ranges from 10 to 15 cm. Since the wheat spikes are found protruding at the upper layer of the canopy, the upper 35 cm of the canopy was clipped out. We used this approximate spike layer clipped from the point cloud for further processing which helped to reduce the computational burden.

## 3.2 SEGMENTATION

The two segmentation techniques short-listed from the literature review are

1. Voxel-based Connected Components Segmentation
2. Mean Shift Segmentation

These two segmentation techniques were adapted to identify and remove the air returns due to multi-edge effect. The ear-classification procedure that was developed as a part of this study was applied on the segmentation output. Finally the ear count from the voxel-based segmentation was compared with that of the mean shift segmentation.

### 3.2.1 Voxel-Based Connected Components Segmentation

The pre-processed points representing the approximate spike layer was used as input. This approach first splits the point cloud into equal sized cubes and continues to eliminate noisy points followed by a connected components segmentation. The following steps were carried out in Matlab to segment the input point cloud as show in Figure 3.2.

(i) **Voxel Definition:** A voxel maybe defined as a 3D pixel with volume associated with it. A voxel coordinate system was defined for the point cloud with $origin(X_v, Y_v, Z_v) = (0, 0, 0)$ and equal sided voxels of size, say $side = 1cm$. The position of each point in the voxel space was predicted using

$$(X_v, Y_v, Z_v) = \left\lceil \frac{(X, Y, Z) * 100}{side} \right\rceil \tag{3.1}$$

$side$ = length of the voxels in cm.
$X, Y, Z$ = the coordinate of the point in metre.
$X_v, Y_v, Z_v$ = the top right corner coordinate in cm of the voxel within which the point is present.

(ii) **Voxel-Based Thinning:** For each voxel, the number of points inside the voxel was counted. For further processing, only the voxels that satisfied the following conditions were used:

- Should have more than one point within the voxel.
- At least one of its direct neighbouring voxels (six direct neighbours which it shares a face) should have more than or equal to a user-defined minimum number of points, say $thinning\_threshold = 2points$.

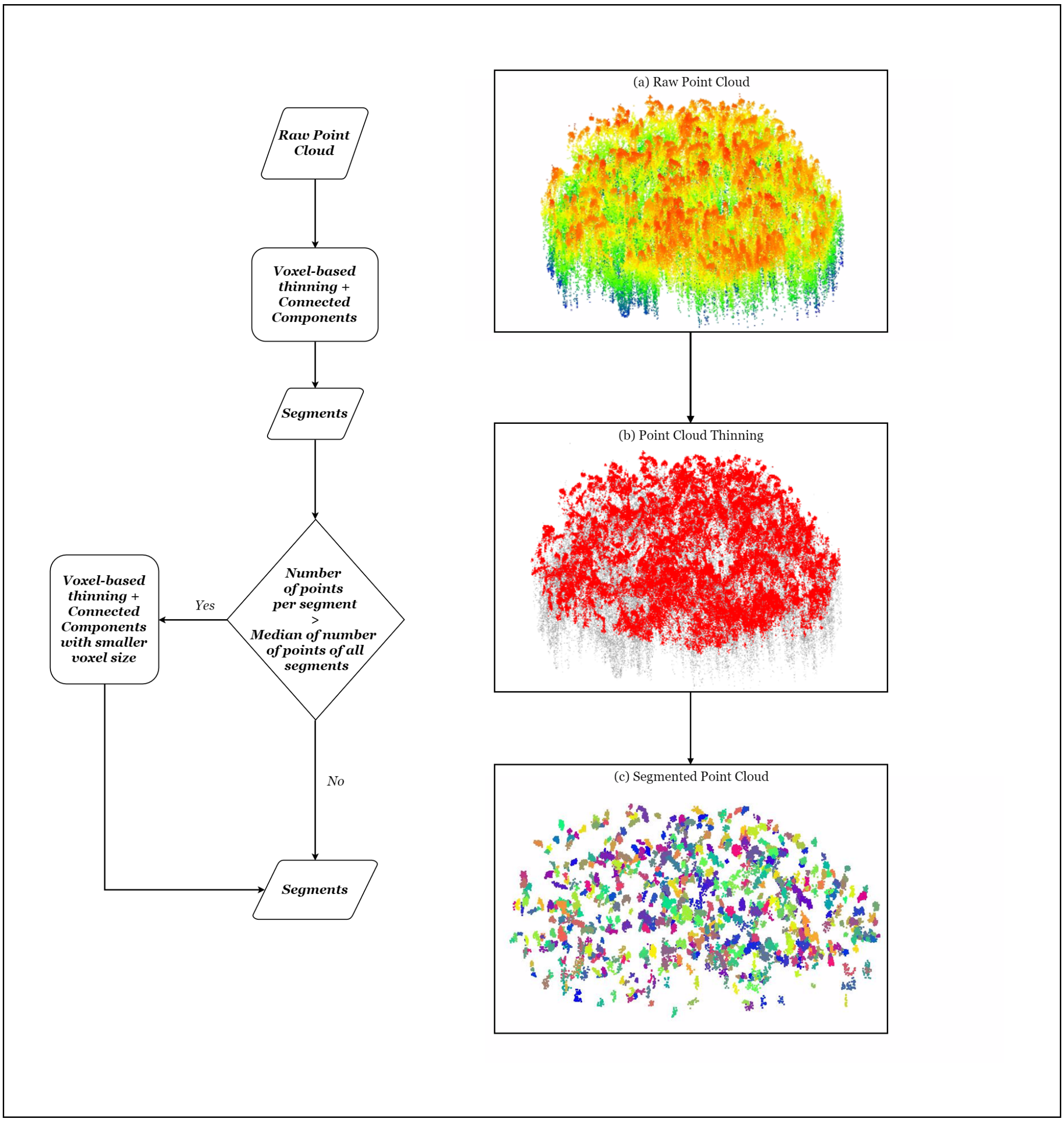



Figure 3.2: Flow-chart depicting the steps involved in the Voxel-based connected components segmentation illustrated with a sample point cloud. (a) Raw point cloud colour coded with height values (b) After point cloud thinning, the retained points are displayed in red and the eliminated points are shown in grey. (c) Final segmentation results with each segment assigned a random colour.

The voxels eliminated based on the above criteria are the ones that contain hovering points and non-ear points with low neighbouring point density. This is illustrated in 3.2 (b) where the eliminated points are shown in grey while the retained points are displayed in red. Hence, the thinning process ensures only voxels that belong to plant organs with high point density are selected for further processing.

(iii) **Connected Components Analysis:** On the voxels selected from the thinning process, a connected components analysis was performed. This will cluster points belonging to a single object and assign a unique segment number to it. The connected voxel components were identified as given below. For each voxel,

- Check if a segment ID has already been assigned to the voxel.
- If not,
  - Check and gather the list of its direct six neighbouring voxels that were retained from the thinning process.
  - If the list is not empty, then for each voxel in the list, search for its direct six neighbouring voxels that were retained during the thinning process.
  - The above two steps are repeated for all voxels in the list until there are no direct neighbouring voxels that were retained from the thinning. Thus, this process is continued until all the connected neighbours are identified.
  - Assign a common segment ID to these connected voxels.
- If yes, move on to the next voxel.

(iv) Based on an initial analysis conducted with manually extracted segments, it was found that a wheat spike contains at least 30 points. Hence, the number of points per segment was calculated and the segments with less than 15 points were removed. The threshold 15 was chosen to consider the point-cloud thinning.

(v) **Second iteration of Thinning and Segmentation:** In areas of high overlap, the objects appear connected and as a result, two or more ears are clustered together as a single segment. In-order to address these cases of under-segmentation, a second iteration of point cloud thinning followed by connected components segmentation is carried out with a smaller voxel size. This process removes the points around the overlapping parts and labels the individual ears as separate segments. The segments for the second iteration of thinning are selected by first calculating the number of points in each segment and finding the median value. If the number of points per segment was higher than the calculated median value, then those segments were assumed to be under-segmented and subject to a second iteration of thinning and segmentation. Steps (i) to (iv) were carried out for those segments with half the voxel size (say, $side\_adapted = side/2 = 0.5cm$) and half the threshold for number of points to be present in the nearest neighbours (say, $thinning\_adapted = thinning\_threshold/2 = 1point$). New segment IDs were assigned.

As more than half of the ear segments are in general correctly segmented in the first iteration, using the median value to choose the segments for the second iteration ensures that the under-segmented ears and leaves are subject to thinning. This assumption might be incorrect in some cases; hence for each segment subject to second iteration of thinning, if the number of points in all the resulting segments fall below the $smallsegment\_threshold$, then the original segment was retained.

The point cloud in Figure 3.2 (c) shows the segments extracted from the raw point cloud by the voxel-based approach. These segments will be labelled as ears and non-ears in the ear classification step. The table 3.2 contains the list of parameters used in the voxel-based segmentation accompanied with a brief description.

Table 3.2: Parameters used in the voxel-based segmentation. Blue rows: User-defined Parameters; White rows: Default Parameters.

| Parameter Name | Description |
|---|---|
| side (*s*) | Length of the side of voxel cubes. |
| thinning_threshold (*tth*) | Threshold for points/voxels removal during voxel-based thinning. (the minimum number of points that should be present in at least one of the direct six neighbouring voxels) |
| smallsegment_threshold (*sst*) | Threshold for removing small segments. (the minimum number of points per segment to be considered as ear segment) |
| median_adapted | Threshold to decide if a segment should be subject to second iteration of thinning and segmentation. (median value of number of points per segment for all segments) |
| side_adapted | Length of the voxel cells for the second iteration of thinning and segmentation. ($X = Y = Z = side/2$) |
| thinning_threshold_adapted | Threshold for the second iteration of voxel-based thinning ($thinning\_adapted = thinning\_threshold/2$) |

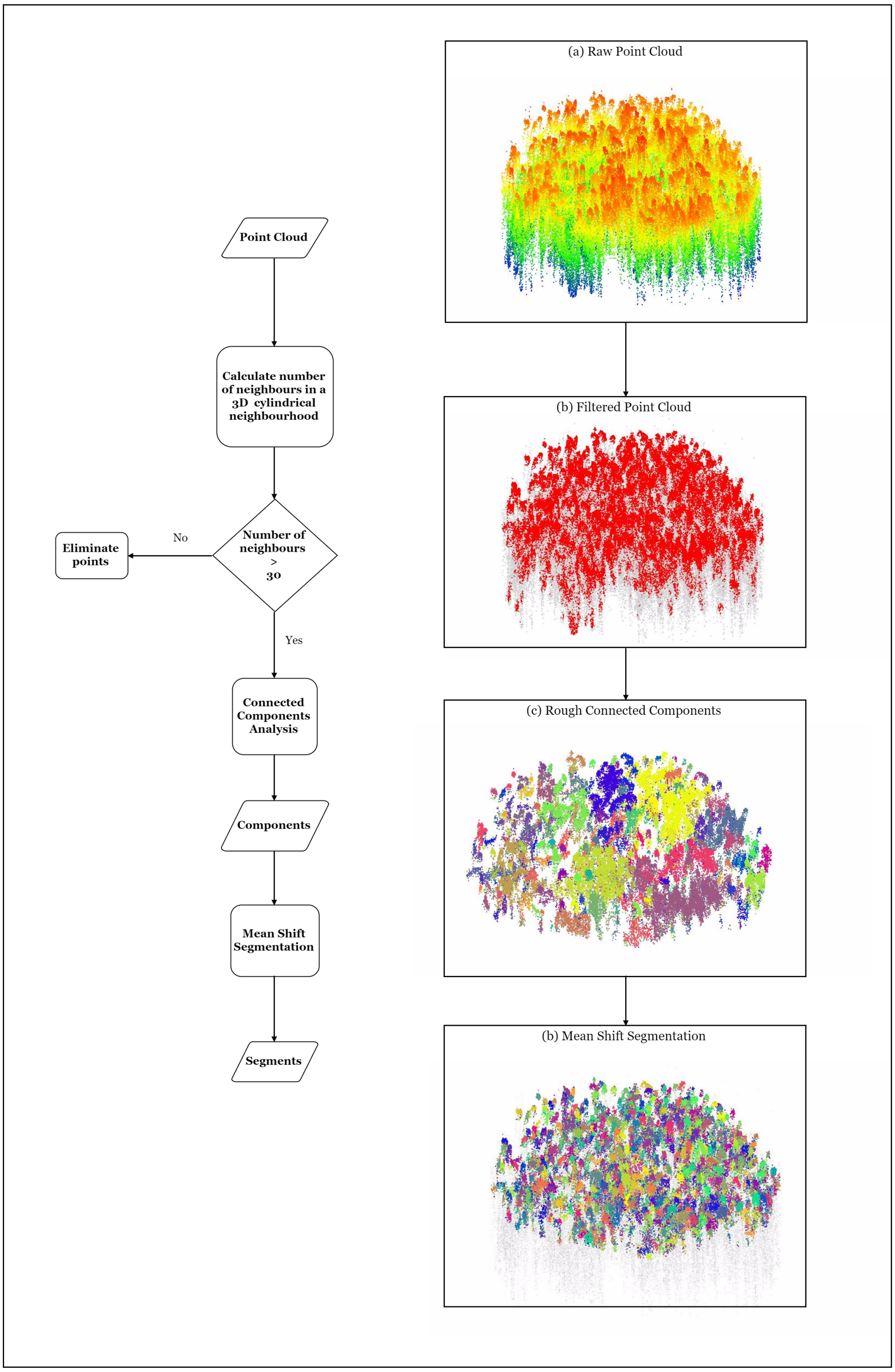
### 3.2.2 Mean-Shift Segmentation

Next, mean shift segmentation method, which is a point based approach is used to design a segmentation strategy to cluster ear segments. Mean shift uses a probability density estimator function to search for modes in the data and cluster the points that fall within function for the user-defined kernel bandwidth. In this study, the three coordinates (X,Y,Z) were used as the kernel bandwidth parameters. The following steps as shown in Figure 3.3 were carried out:

(a) **Filtering:** In order to remove the points hovering over the canopy and recorded from the stem region, the following steps were carried out:

- With each point as centre, the number of neighbours in a 3D cylindrical neighbourhood ($X = 3cm, Y = 3cm, Z = 6cm$) was calculated.
- Only points with more than 30 neighbours in this cylindrical neighbourhood would be used in further processing.

The dimension of the cylindrical neighbourhood was decided based on a trial and error method which helped to identify the ideal dimensions and the range of neighbouring density within which wheat ears could be approximately identified (Refer to Appendix A). The threshold 30 was decided based on an initial analysis conducted with manually extracted segments, which showed that a wheat spike contains at least 30 points. Figure 3.3 (b) shows the points eliminated in this step in grey and the points retained in red.




.

Figure 3.3: Work-flow involved in Mean Shift Segmentation illustrated with a sample point cloud (a) Raw point clouded colour coded with height values (b) Points filtered based on the point density in their cylindrical neighbourhood. Grey points had less than 30 neighbours and hence were removed whereas the red points had more than 30 neighbours and so were retained. (c) Connected components analysis to identify under-segmented ear blobs (d) Mean shift segmentation results displayed by assigning random colours to the segments

(b) **Connected Components:** An initial connected components analysis was performed on the point cloud to identify under-segmented blobs of connected points as can be seen in Figure 3.3 (c). These blobs contain a few ears and the next step aims to identify the individual ears in the blobs using mean shift segmentation. This rough connected components step based on proximity and neighbourhood definitions was included to reduce the processing time.

(c) **Segmentation** For each component, mean shift segmentation was carried out using an executable from the Mapping Library developed at ITC, Enschede. Only the (X, Y, Z) coordinates of the points were used as input parameters in the algorithm. Hence the mean shift segmentation identifies clusters in the 3D space that gives a minimum gradient value calculated from the kernel function for the kernel bandwidth defined by the user. The kernel bandwidths were selected as $X = Y < Z$ following a prolate spheroid, to approximate the average size and shape of a wheat ear. The segments identified by the algorithm are displyed in Figure 3.3 (d) in different colours.

(d) Based on an initial analysis conducted with manually extracted segments, it was found that a wheat spike contains at least 30 points. Hence, the number of points per segment was calculated and the segments with less than 20 points were removed. The threshold 20 was decided to take into consideration the initial points removal carried out as a part of the filtering process

Initially, segmentation experiments were carried out by applying the mean shift method directly on the entire point cloud. But the processing time was quite high for even a small subset of the point cloud. This was because, for dense point cloud, the shift step is smaller as the mean shift vector moves towards areas of higher point density. Hence, convergence of the multiple clusters near the local maximum is quite slow. Thus, to reduce the computation time, it was decided to first perform a rough connected component analysis followed by the mean-shift segmentation on the resulting components. This is helpful due to the comparatively lesser number of points in the components which ensures that the convergence occurs faster. The list of parameters used during the mean shift segmentation is shown in Table 3.3

Table 3.3: Parameters used in the Mean Shift Segmentation. Blue rows: User-defined parameters.

| Parameter Name | Description |
|---|---|
| Kernel Bandwidth $(X, Y, Z)$ | Parameters needed to estimate the mean shift vector for clustering in X, Y and Z dimensions |
| smallsegment_threshold (*sst*) | Threshold for removing small segments. (the minimum number of points per segment to be considered as ear segment) |

#### 3.2.3 Ear Classification

After segmenting the point cloud using either the mean-shift method or the voxel-based method, we are left with unlabelled segments that could be wheat ear, leaf, stem or other plant parts. Hence, a methodology was developed (Figure 3.4) to distinguish the ear segments from the output segments and extract them.

1. **Top-Most Segment Selection:** Since the wheat ear is present at the top of the canopy, the first step was to extract the top-most segments in the segment-space. This was done as follows:

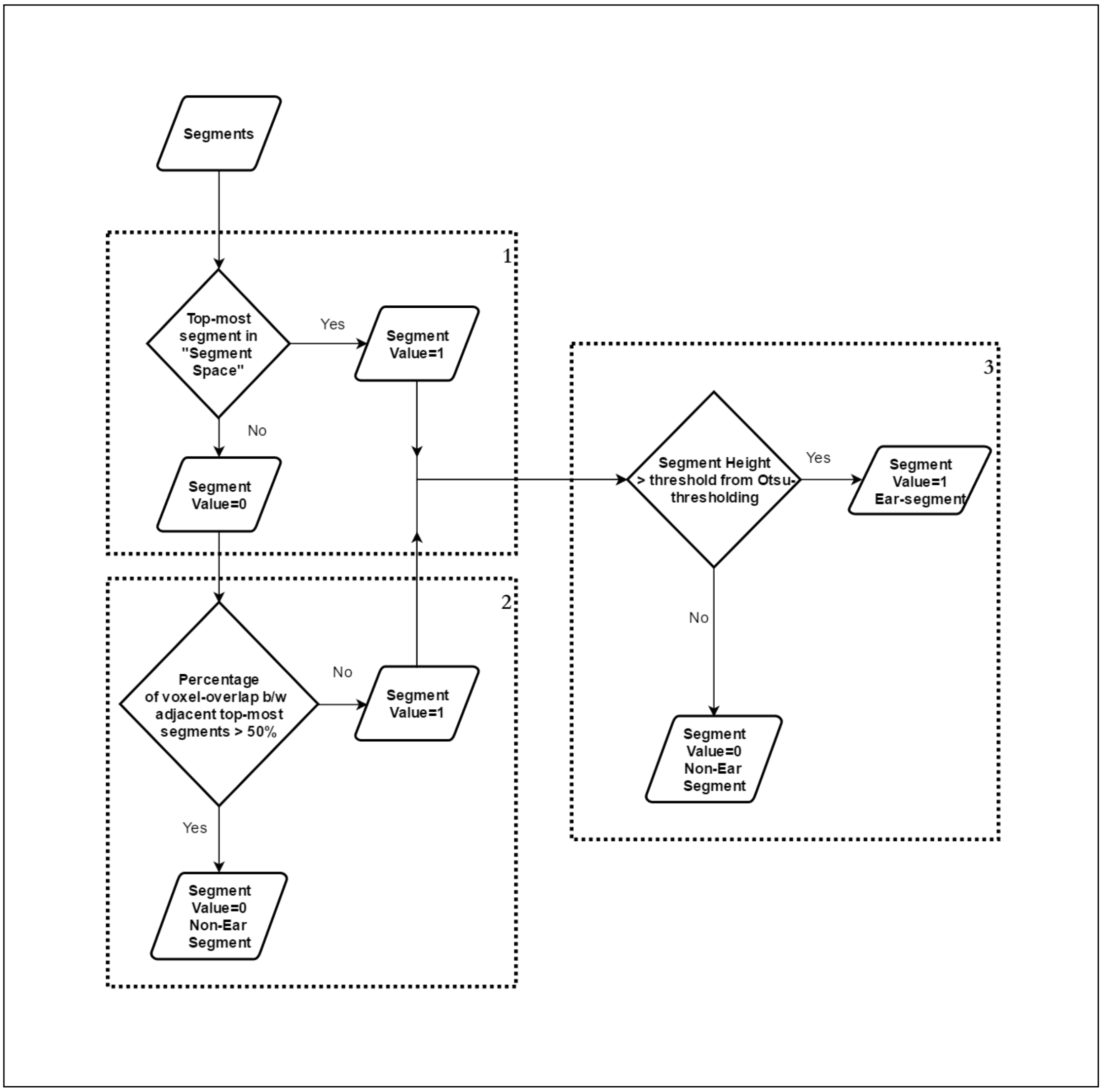



Figure 3.4: Strategy proposed to classify segments into ear and non-ear segments

- Find the highest point per segment and sort the segments by this maximum height, in descending order.
- Loop through the sorted segments one by one.
- For the first segment in the sorted list, the segments present below this top-most segment was found and temporarily labelled as non-ear plant parts and removed from the segments list. Also, label the top-most segment as ear.
- Move on to the next segment with highest point in the sorted list and execute the previous step.
- Carry on until the all the segments in the segment list have been labelled.

Now, the temporarily labelled non-ear segments are reconsidered where their extent of overlap with the top-most segment is evaluated. If the overlap percentage is above a certain value, say 50%, then the segment is assumed to be directly below an ear segment and permanently labelled as a non-ear segment. In some scenarios where the ears are bent and overlapping, as highlighted in the green box in Figure 3.5(a), the ear segments from the shorter plant initially tend to get mislabelled as non-ears. However, reconsidering these segments based on the extent of their overlap with the top-most segments ensures that they are correctly relabelled as ear segments as shown in the green box in Figure 3.5 (b).

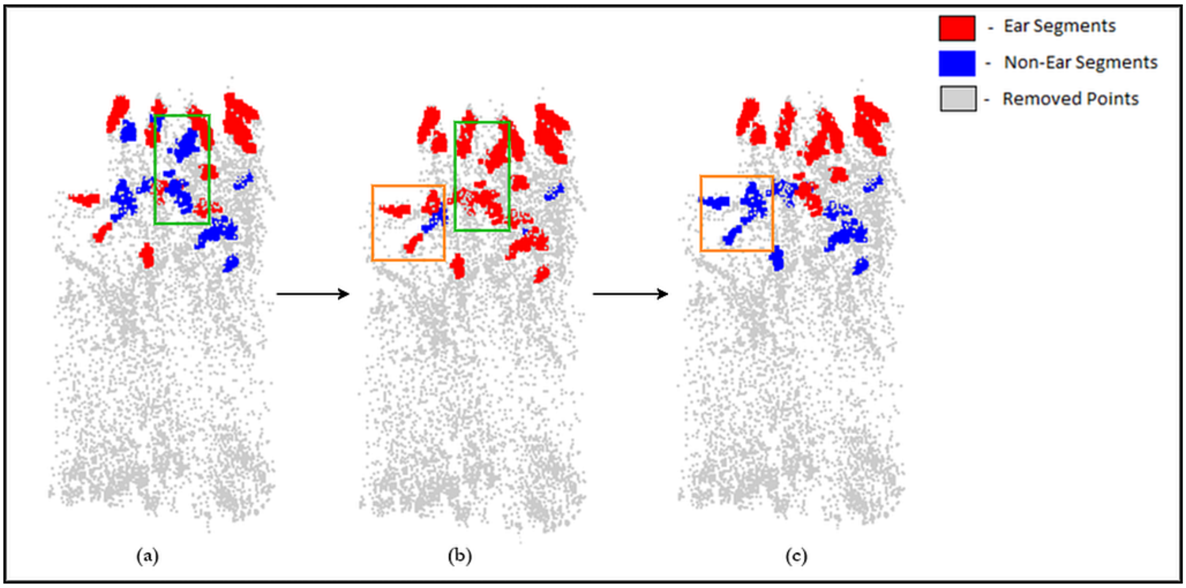



Figure 3.5: Example to illustrate the steps involved in Ear classification. Ear segments are displayed in red while the non-ear segments are displayed in blue (a) Ear segments identified by searching for the top-most segment in segment space. Green box highlights a case of mislabelling due to overlap (b) Relabelling non-ear segments depending on extent of overlap with the segments above. Orange box focuses on a case of mislabelled non-ear segments. (c) Final segment labels assigned based on height thresholding selected by Otsu's method which identifies the trough between two distinct peaks in a histogram.

2. **Height Thresholding:** After the above steps, some of the non-ear segments might still be wrongly labelled as ear (an example as shown in the orange box in Figure 3.5 (b)). On analysing the height histogram (Figure 3.6) of the segments labelled as ears, two distinct peaks could be identified; one peak corresponding to the ear segments and the other one to the non-ear segments. This is due to the fact that there are always more number of points describing the ear segments since they are present on the top of the canopy. Hence, in most cases, they are not overshadowed by other plant parts. Hence, identifying the trough between the two peaks of the height histogram will aid in correct labelling of the ear segments. This height threshold is selected using Otsu's threshold selection method (Otsu, 1979) for histograms. The segments that lie below this height threshold but were classified as ear in the previous steps are reclassified as non-ear segments. No changes are made to the labels of the segments above this height threshold. The orange boxes in figure Figure 3.5 (b, c) highlight a mislabelled ear segment being corrected by the thresholding step.

Thus, the grammar for classifying the ear segments is based on the geometry of the wheat ears and the canopy characteristics. The parameters used during the process are briefed in Table 3.4.

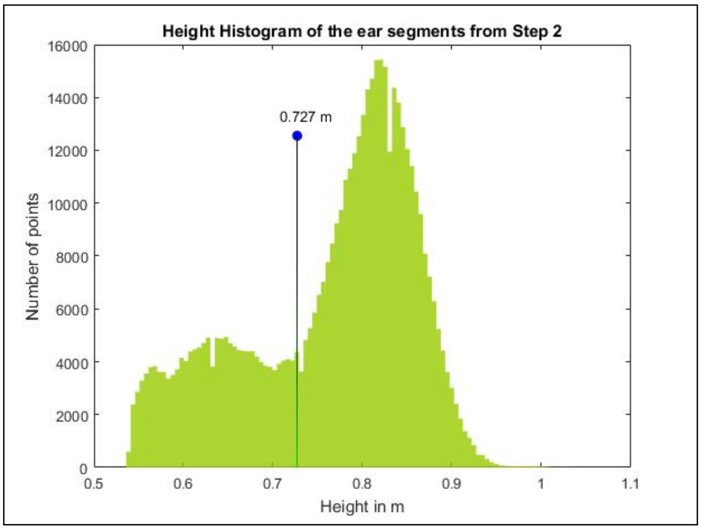



Figure 3.6: Height histogram showing the two distinct peaks between the ear and non-ear segments. The line at *0.727 m* denotes the threshold chosen from Otsu's thresholding.

Table 3.4: Parameters used in classification of the segments. Blue rows: User-defined parameters; White rows: Default Parameters.

| Parameter Name | Description |
|---|---|
| overlap_percentage (*op*) | Percentage of overlap between voxels of a non-ear segment(from Step 1) and the ear segments present above it. |
| height_threshold | The low point between the two peaks present in the height histogram of all ear segments identified till Step 2.The two peak correspond to ear and non-ear plant parts and the threshold is selected by Otsu's thresholding method. |

The methodology was designed keeping in mind the geometry and orientation of the wheat ears that change with the flowering stage of the plant. It also addresses the problem of detecting the ears of stunted plants present in the field that might be partially hidden by neighbouring taller plants. The performance of the ear classification in combination with the two short-listed segmentation techniques is analysed in detail in the forthcoming chapters.

# Chapter 4
# Parameter Optimization

The ear-detection methodology developed in the previous chapter uses a combination of parameters to get the final ear count. This chapter gives a description of how the input parameters were optimised and the sensitivity of the results to their values. A sensitivity analysis of the ear detection techniques was conducted to understand the influence of the different parameters on the final ear count. This sensitivity test helped to identify which among the user-defined parameters are the most influential. This was followed by a 3-fold cross validation to optimise the values for input parameters.

## 4.1 SENSITIVITY ANALYSIS

The Tables 3.2 through 3.4 show the user-defined and default parameters used during the segmentation and ear-classification process. Sensitivity analysis was performed by running the ear-detection process for different values of the input parameters. This was done by varying one parameter value while keeping the others constant; this was repeated for all the parameters. Analysing these results helps to quickly identify the input parameters that are comparatively more influential on the final results. Hence, for our study, it was decided to conduct two sets of sensitivity analysis

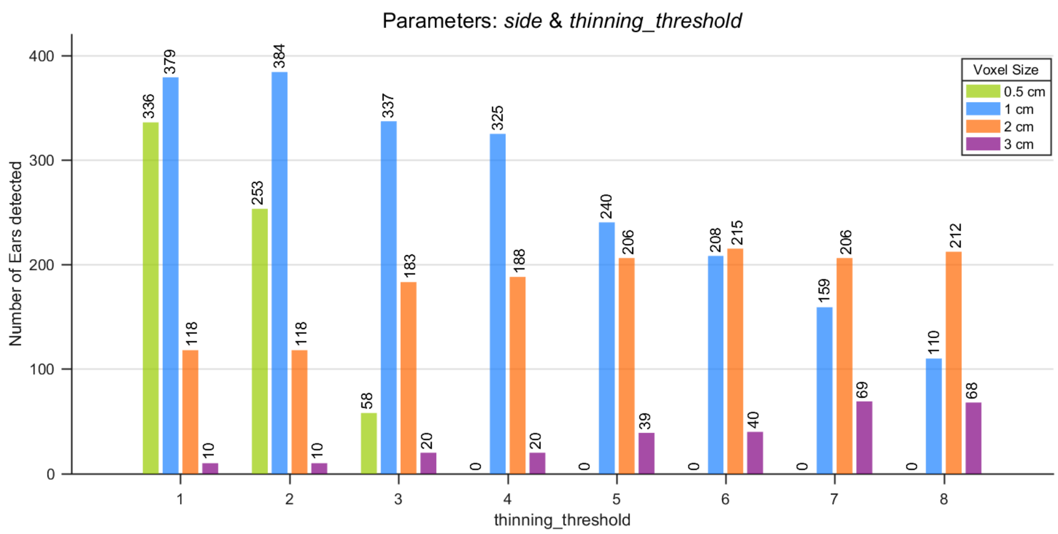
- For the input parameters used in the voxel-based ear detection approach
- For the input parameters used in the mean-shift based ear detection approach

These tests were conducted on a small subset from *Plot A*, with an approximate area of $1m^2$, acquired on 10 June 2016 (refer to Table 3.1).

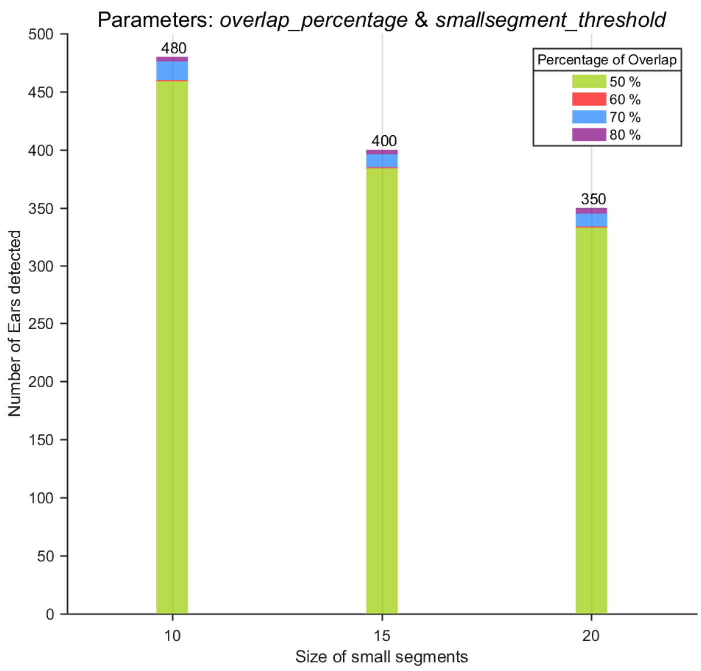
### 4.1.1 Voxel-Based Ear Detection

In the first step, the input parameters $side$ and $thinning\_threshold$ were varied while keeping the other parameters constant. The number of points within a voxel depends on the size of the voxels used; bigger the size of the voxel, larger the number of points within it. Moreover, the size of the voxels should be fixed depending on the average distance between the crops whereas the threshold for voxel removal should be decided based on the voxel size and point density of the dataset. The Figure 4.1 (a) shows the change in the number of ears detected for different combinations of the input parameters $side$ and $thinning\_threshold$.

The graph supports the fact that the ear detection rate depends on the size of the voxels in combination with the thinning threshold. When voxels of $side = 0.5cm$ are used, the number of points falling within each voxel is less. As a result, using a strict threshold of higher than 3 points for voxel removal results in over-thinning of the point cloud. (shown with green bar in Figure 4.1 (a)). Conversely, when bigger voxels are used with a lenient threshold for points removal, it results in under-segmentation. On conducting a qualitative analysis of the results, it was found that for voxel size 3, there were severe cases of under-segmentation in several regions of the plot, especially in densely cropped areas. This is also reflected in the Figure 4.1 (a) by a sharp decrease in the number of ears detected.Thus, it can be inferred from the graph that both these input parameters are crucial to get acceptable segmentation results.




(a) The voxel size and threshold for voxel removal are found to be interdependent and influential




(b) The threshold for removing small segments and percentage of overlap between voxels are comparatively less influential.

Figure 4.1: Sensitivity of Ear detection to the input parameters of Voxel-Based Segmentation. (*Plot A*, 10 June 2016)

The Figure 4.1 (b) shows the influence of the $smallsegment\_threshold$ and $overlap\_percentage$

on the final counts. We can infer from the graphs that the role of the parameter $overlap\_percentage$ is not very influential in the final results. Also, the number of ears detected was constant for values above 80%. However, in order to ensure that leaf segments that are directly below an ear segment are labelled appropriately, it is necessary to have a value of 50% or higher. This also ensures that the partially overshadowed ear segments are not misclassified as non-ear segments. After segmentation, small segments with less than $smallsegment\_threshold$ number of points are removed. This threshold once again depends on the point density of the dataset.

#### 4.1.2 Mean Shift-Based Ear Detection

The only user-defined parameter used during the mean shift segmentation is the kernel bandwidth. The sensitivity of the ear counts to different combinations of the $(X, Y, Z)$ kernel bandwidth parameters were tested. These combinations were designed such that $X = Y < Z$, taking the approximate shape of the wheat ear into consideration. Hence, values for $X, Y$ were varied only upto 2 cm, the average diameter of the wheat ear is close to that range. The value for $Z$ coordinate bandwidth was varied from 2 to 6 cm which is in accordance to the average height of the wheat ear.

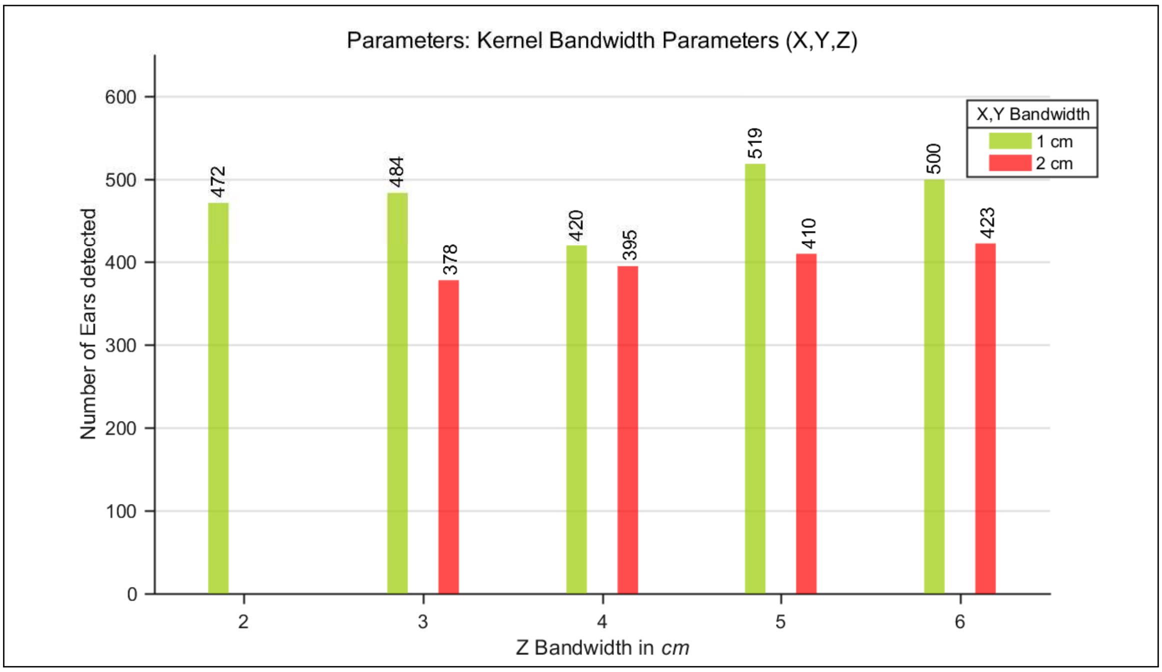



Figure 4.2: Sensitivity of the Ear detection rate to Mean Shift Segmentation Parameters i.e. Kernel Bandwidth Parameter: X,Y,Z (*Plot A*, 10 June 2016).

In Figure 4.2, it is noticed that the sequential increase in bandwidth size does not produce a predictable trend in the number of ears detected. The different combinations of $(X, Y, Z)$ values result in the formation of different clusters in the 3D space. For example, using a higher $Z$ coordinate bandwidth results in undersegmented plant parts of points from the ear and stem regions. This factor along with a qualitative analysis is explained in detail in Subsection 4.2.2. The other two user-defined parameters for the mean shift-based ear detection are the threshold for the removal of small segments and the threshold for allowable extent of overlap between two ear segments. The sensitivity of these two thresholds on the ear count has already been presented in the previous section (refer to Figure 4.1 (b)).

### 4.2 PARAMETER TUNING

In order to optimise the values for input parameters, a combination of qualitative and quantitative analysis was used. For each input parameter, the range of values that produces reasonable segmentation results was identified by qualitative analysis of the segments. An optimal value from this range was chosen by a 3-fold cross validation. Figure 4.3 shows the experimental set-up designed to carry out the cross validation.

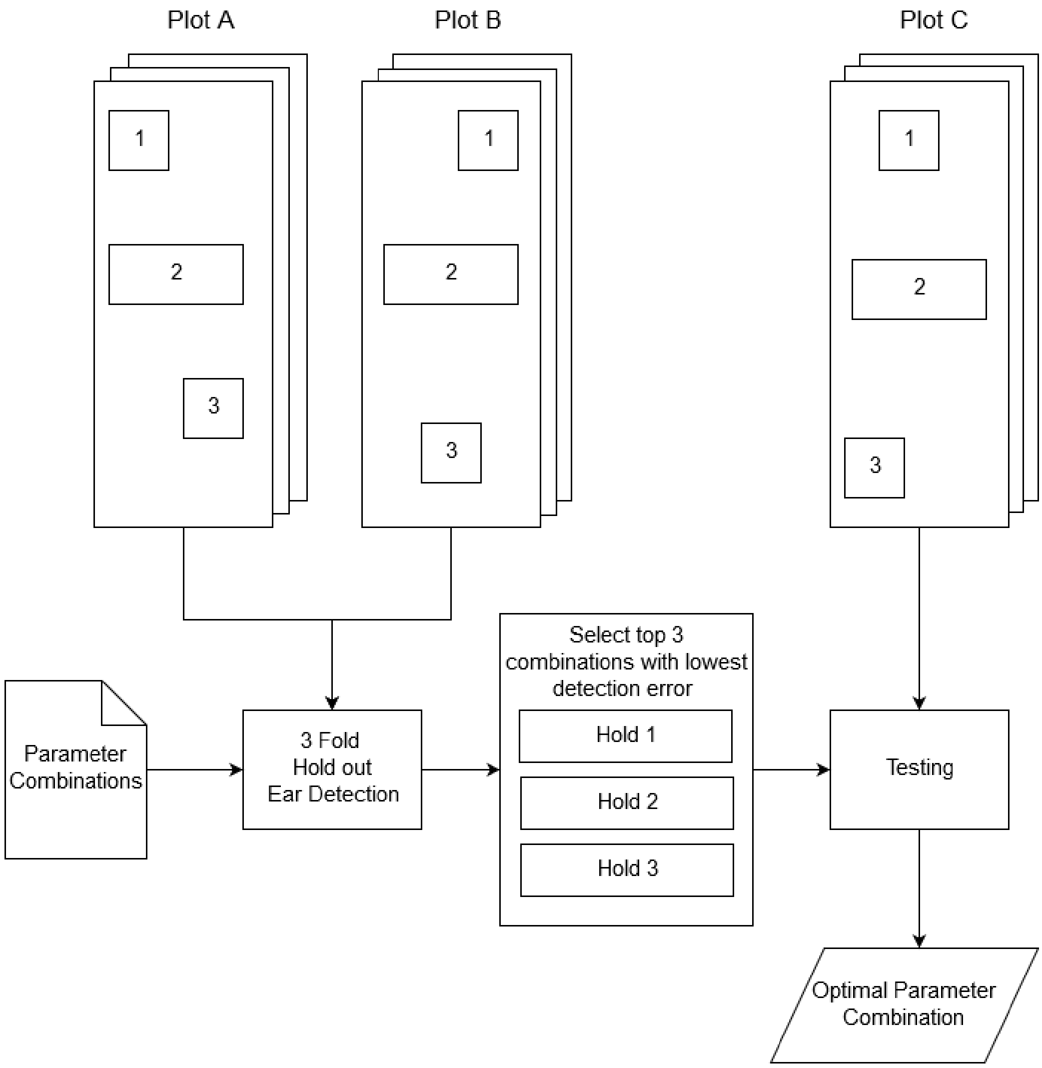



Figure 4.3: Experimental set-up used in the 3 Fold Hold Out Cross Validation for Parameter Tuning

For the purpose of parameter tuning, three subsets were prepared from each dataset acquired on the three dates i.e. nine subsets for each of the three plots. Among these, the subsets from two plots, *Plot A* and *Plot B* were used for tuning and the subsets from *Plot C* were used for validation. Thus, the parameters were tuned based on two varieties of wheat from different developmental stages and validated against a wheat variety that was not used during the tuning process. The steps involved in parameter tuning were as follows:

(i) **Tuning:** For each fold, two subsets from each dataset of *Plots A* and *Plot B* were used to evaluate the detection rate against the manual counting from the field. The combination of parameters with the highest detection rate was chosen for each fold.

(ii) **Validation:** The three combination of parameters chosen from the tuning were then tested on the subsets prepared from *Plot C*. The parameter combination that gives segments of reasonable quality along with a high detection rate was chosen after carrying out a qualitative analysis of the ears detected.

#### 4.2.1 Voxel-Based Ear Detection

For each of the four user-defined parameters, a range of values was chosen to be tuned and evaluated on the dataset.

(i) ***Side*** **:** The size of the voxels should be fixed according to the average spacing between the plants in the field. Wheat crops, in general, are not cropped with constant spacing. Hence, it is common to find varying crop densities throughout the field. In our study, there was irregular spacing among the crops varying between $0.5cm$ to $3cm$ with an average point spacing of $16pts/cm^2$. Thus, the size of the voxels was varied from 0.5 to 2 $cm$.

(ii) ***Thinning _threshold*** **:** The values for thinning threshold was used depending on the size of the voxels. For smaller voxel size, a lower threshold was tested, likewise for bigger voxels, a higher threshold was used.

Table 4.1: Optimal Values for the input parameters of Voxel-Based Ear Detection.

| Parameter Name | Optimal Value | Average Error |
|---|---|---|
| side | $1cm$ | RMSE: 41.43 ears/$m^2$ MAPE: 15.43 % |
| thinning_threshold | 2 points | |
| smallsegment_threshold | 15 points | |
| overlap_percentage | 70% | |

(iii) ***smallsegment_threshold*** **:** For the datasets used in this study, the average number of points per ear was always found to be higher than 30. Hence, segments with less than 30 points could be considered as non-ear segments. However, a threshold ranging from 10 to 20 was used so as to consider the thinning of the point cloud.

(iv) ***overlap_percentage*** **:** As discussed in Section 4.1.1, the threshold for allowable overlap between two ear segments has comparatively minimum influence on the detection rate. However, a range of 50% to 80% was used in the tuning process.

The Table 4.1 lists the optimal input parameter values for situations comparable to the ones in this study. These were chosen by qualitatively comparing the best performing results, i.e. minimum error in ear density estimates, from each fold of the three fold hold out cross validation technique. The average error values shown in the table were calculated by taking the mean of the RMSE and MAPE values obtained in the validation step of the three fold hold out method.

#### 4.2.2 Mean Shift-Based Ear Detection

By performing the 3 fold hold out method, certain combinations of parameters were identified to have good detection rate in terms of the number of ear segments detected. However, a visual examination of the results revealed that the segmentation performance was not uniform throughout the point cloud. There were cases of poor segmentation in different regions of the point cloud.Thus, the optimal parameters for the kernel bandwidths were chosen based on a qualitative analysis of the segmentation results, combined with their detection efficiency. Table 4.2 shows the optimal parameters identified for the mean-shift based ear detection.

Table 4.2: Optimal Values for the input parameters of Mean Shift-Based Ear Detection.

| Parameter Name | Optimal Value | Average Error |
|---|---|---|
| Kernel Bandwidth: X | $1cm$ | RMSE: 69.43 ears/$m^2$ MAPE: 20.43 % |
| Kernel Bandwidth: Y | $1cm$ | |
| Kernel Bandwidth: Z | $2cm$ | |
| smallsegment_threshold | 20 points | |
| overlap_percentage | 70% | |

#### 4.2.3 Tuning over varying Point Density

From the above subsection, it was understood that the point density of the LiDAR observations play an important role in the decision of the input parameter values. Thus, it was decided to downsample the point cloud by retaining 75%, 50% and 25% of the original points (The point density of the original dataset is $16 points/cm^2$). The optimal parameters adapted to suit the downsampled point clouds was identified for both the voxel-based and mean shift segmentation methods. The values of parameters that are not dependent on the point density such as the *X,Y,Z* kernel bandwidth parameters, *overlap_percentage (op)* and *side (s)* were not changed and were assigned values as shown in Tables 4.2 and 4.1. The other input parameters used in the implementation were assigned values as listed in Table 4.3.

Table 4.3: The different Input Parameters used according to the Point Density of the datasets.

| Percentage of Points Retained | Voxel-Based | | Mean-Shift |
|---|---|---|---|
| | *tt* | *sst* | *sst* |
| 25 % | 1 | 10 | 13 |
| 50 % | 1 | 13 | 15 |
| 75 % | 1 | 18 | 18 |
| Original | 2 | 15 | 20 |

In Table 4.3, it is noted that when the point density is higher, a comparatively higher $thinninig\_threshold(tt)$ is identified to be optimal. Intuitively, when a lenient $thinninig\_threshold(tt)$ is used for the removal of noisy points, the threshold identified for the removal of small segments is higher.

Chapter 5

# Performance Evaluation

The robustness of the ear detection methods developed in this study are evaluated against three different criteria. In order to evaluate the performance of the developed ear-detection approaches, subsets of approximately $1m^2$ were prepared from the 9 different datasets available (as shown in Table 3.1). The ear-detection procedures were then implemented on these subsets and the results were compared with manual counts from the field. The description of the criteria and experiments used in the evaluation process and the corresponding results are elucidated in the subsequent sections.

## 5.1 EVALUATION METRICS

This section describes the metrics used in this chapter to evaluate the performance of the ear detection methods. A combination of quantitative and qualitative metrics was used throughout the evaluation process.

### 5.1.1 Root Mean Square Error (RMSE)

The Root Mean Square Error was used as one of the measures to evaluate the error associated with the detection of number of wheat ears per plot. The performance of the algorithm might be different for different varieties, flowering stage and might also be influenced by factors like cropping density. Thus, the ear detection method being evaluated might produce good estimates for a particular plot while returning poor estimates for another one. Therefore, *RMSE* was chosen to represent the collective error associated with the ear count over different plots and plant developmental stages as it is sensitive to outliers (Chai & Draxler, 2014).

$$RMSE = \sqrt{\frac{\sum (Ear_{\text{det}} - Ear_{\text{ref}})^2}{N}} \tag{5.1}$$

where,

$Ear_{\text{det}}$ = Number of ears detected by the automatic method.
$Ear_{\text{ref}}$ = Number of ears identified by manual counting.
$N$ = Number of datasets used in the evaluation.

From Equation 5.1, we infer that the magnitude of individual errors have a direct influence on the *RMSE* value calculated.

### 5.1.2 Mean Absolute Percentage Error (MAPE)

The Mean Absolute Percentage Error gives a measure of the overall performance of the ear detection algorithm by providing an averaged absolute precision in terms of percentage. The *MAPE* provides an estimate that is independent of the spatial extent of the plot. Hence, using the *MAPE* in combination with $RMSE$ gives an overall understanding of the performance of the algorithm across different plots.

$$MAPE = 100\frac{1}{N}\sum\left|\frac{(Ear_{\text{det}} - Ear_{\text{ref}})}{Ear_{\text{ref}}}\right| \tag{5.2}$$

where,

$Ear_{\text{det}}$ = Number of ears detected by the automatic method.
$Ear_{\text{ref}}$ = Number of ears identified by manual counting.
$N$ = Number of datasets used in the evaluation.

### 5.1.3 Qualitative metrics

Besides the quantitative metrics explained above, qualitative analysis of the ear segments was performed by visually examining the extracted ear segments. Are there situations of severe under-segmentation or over-segmentation? Are there missing segments because of the point cloud thinning or removal of small segments? Is there predominant misclassification between the ear segments and non-ear segments? Also, the efficiency of the two methods in terms of processing time was analysed. Since different programming languages and libraries were used to execute the two methods, a quantitative analysis of the processing time would be complicated. Hence, a simple qualitative analysis was done.

## 5.2 EVALUATION CRITERIA

The following three criteria were used to understand the robustness of the two ear detection methods developed in this study.

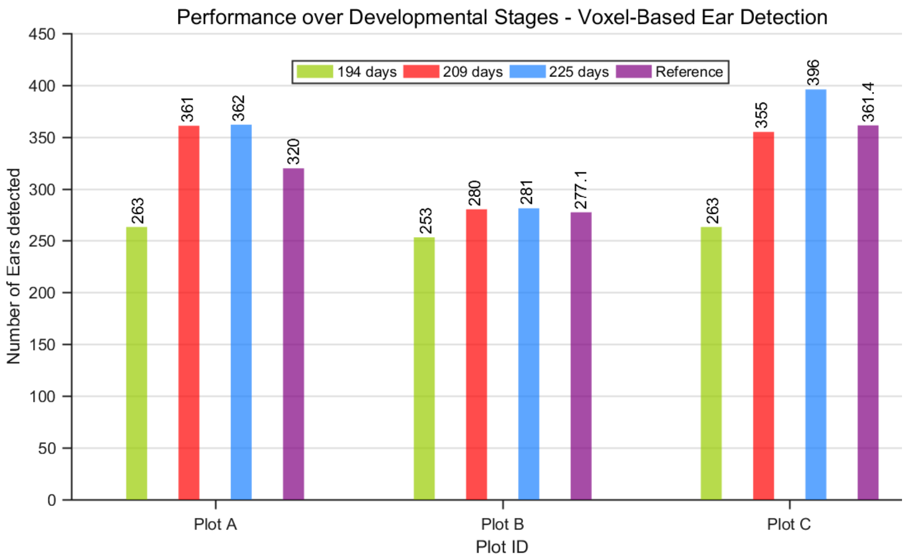
- **Crop developmental stage:** The size and orientation of the wheat ear changes throughout its flowering stage(refer to Figure 2.1). When the ear is young, it is erect and smaller in size whereas towards the harvest period, it becomes bigger and bent. This trait of the ear growth has an influence on the point density of the ear segments. Thus, using the datasets available for three different developmental stages of the same plot, the effect of crop growth on the ear detection rate was evaluated.

- **Wheat variety and irrigation treatment:** The variety of wheat and water stress are found to influence the size of the wheat grains (Sionit, Hellmers, & Strain, 1980). Hence, the ear detection methods developed were tested on three different varieties of wheat subject to different irrigation treatments. This helped to understand how the methods perform for plots with varying plant density (refer to Table 3.1 which shows the plant density per $m^2$ for each plot).

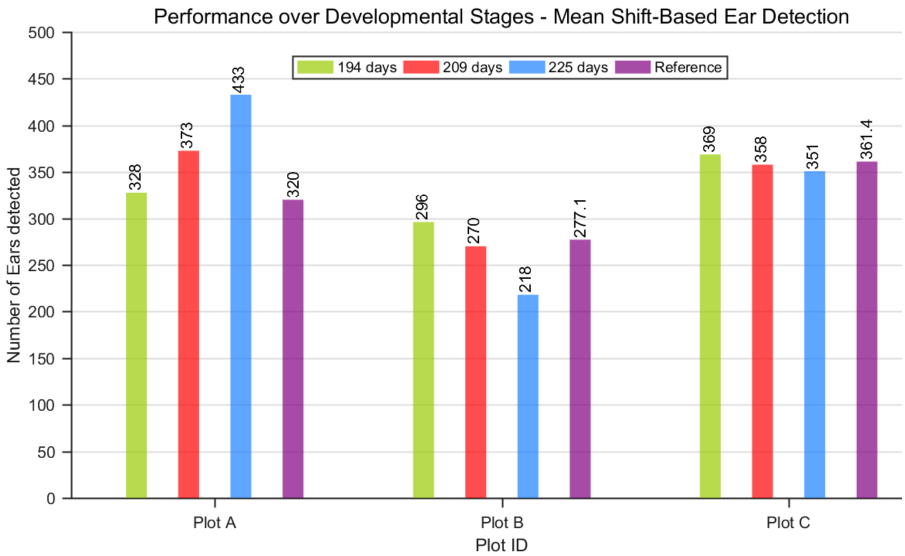
- **Effect of point density:** During parameter optimisation, it was observed that point density is one of the important factors that influences the choice of input parameter values and ear detection rate. Hence, to understand how the ear detection methods perform for different point densities, the point cloud was downsampled. Three downsampled versions of the datasets was created by retaining 75%, 50% and 25% of the original points in a random manner.

### 5.2.1 Effect of Developmental Stage

The ability of the proposed ear detection techniques to handle the temporal variation of the wheat crop was tested. Since the number of plants in a plot does not vary over time, an ideal algorithm is expected to identify the same number of ears throughout the growing season. However, this was not the case as can be observed from Figure 5.1.




(a) The performance of the voxel-based detection method is consistent for crops of age 209 days and above. For younger crops of age 194 days, there is an under estimation in the number of ear segments.




(b) The performance of mean shift-based detection method does not follow any particular trend. On an average, the detection rate is better for the crops of age 209 days and below.

Figure 5.1: Effect of crop developmental stage on the automatic ear detection methods.

For the voxel-based ear detection, an increase in the number of ears identified with increase in the crop age is noticed (Figure 5.1 (a)). This is because, in the early developmental stage, the wheat ears are small in size and described by lesser number of points in the point cloud. In Section 4.2.1, it was explained how the voxel-based approach is heavily dependent on the point density of the point cloud. This explains the noticeable under-estimation in the number of ears for the datasets of young plants of age 194 days. This drawback for younger plants can be overcome by using a lenient threshold during the **"Point Cloud Thinning"** step described in Subsection 3.2.1. For the later growth stages, even though there was comparatively smaller difference between the number of ears detected, cases of over-segmentation was observed.

On the other hand, for the mean shift-based ear detection a mixed trend is observed. In Figure 5.1 (b), a decrease in the number of ear segments identified is noticed for the late developmental stages of *Plot B* and *Plot C*. But for *Plot A*, a trend similar to that of voxel-based approach is noticed. These conflicting trends is due to the following two reasons:

- **Over-Segmentation in crops of age 225 days:** Over-segmentation issues cause an increase in the number of ear segments reported. The clustering of an ear into a single segment depends on the kernel bandwidth. For the late developmental stage, when the size of the ear is big, comparatively large bandwidth size is preferred to avoid over-segmentation issues. This can be overcome by using an adaptive kernel bandwidth for the mean shift segmentation, chosen based on the point density (Ferraz et al., 2010).

- **Misclassification of non-ear segments in crops of age 209 days:** On analysing the classified point cloud, it was noticed that there were some non-ear segments misclassified as ears. The cause for this discrepancy in classification is the ***"Height Thresholding"*** step explained in Section 3.2.3. In order to make a successful ear classification using Otsu's thresholding, it is ideal to have a large number of points from the ears and leaves and comparatively less number of points from the stem region, leading to two distinct peaks in the height histogram. However, unlike the voxel-based ear detection, the mean shift approach does not remove the points on the stem. Also, due to the smaller ear size in younger plants, the number of points on the ear is also less. This results in a single peak in the height histogram rather than the ideal two distinct peaks, causing the misclassification.

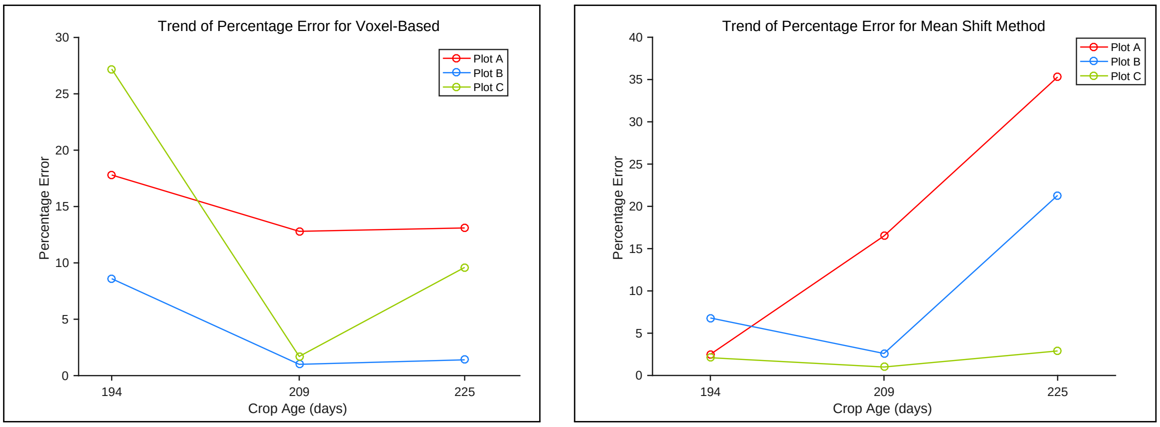
Table 5.1: Evaluation of the effect of crop developemental stage on ear detection.

| **Developmental Stage (Crop Age)** | Voxel-Based | | Mean-Shift | |
|---|---|---|---|---|
| | RMSE ($ears/m^2$) | MAPE(%) | RMSE ($ears/m^2$) | MAPE (%) |
| Early (194 days) | 67.11 | 17.91 | 12.63 | 3.80 |
| Mid-Growth (209 days) | 24.01 | 5.20 | 30.93 | 6.68 |
| Harvest (225 days) | 31.49 | 8.03 | 73.86 | 19.83 |

The performance of the ear detection approaches for the three developmental stages, with the errors aggregated over different varieties, is listed in Table 5.1. From the table, we infer that the voxel-based method performs better for the crops aged 209 days and more with an average percentage error of 5.2% and 8.03%. Whereas, the mean shift-based approach shows relatively better results for the crops of age 209 days and less, with an average percentage error of 3.8% and 6.68%. Thus, it is concluded that the crop developmental stage influences the performance of the ear detection methods. If the age of the plants is known, then the input parameters should be set accordingly for optimal performance.

### 5.2.2 Effect of Wheat Variety and Treatment

From the discussion in the previous section, we infer that the developmental stage influences the segmentation and as a result the ear detection process. Hence, to understand the effect of plant variety on the ear detection methods, the assessment should be independent of the crop age. It was decided to compare their performance for the three different varieties using the datasets acquired during their mid-developmental stage (crops of age 209 days), for which both the ear-detection approaches had comparatively lower errors (as seen in Table 5.1).




(a) Error trends for Voxel-Based Detection. (b) Error trends for Mean Shift-Based Detection.

Figure 5.2: Effect of wheat variety on ear detection represented by Percentage Error. *Plot C* has a higher error in most of the situations.

From Table 5.2, lists the *Absolute Error* and *Percentage Error* in ear detection when tested on crops of age 209 days, for three varieties. It is understood that the error in ear detection for this developmental stage is comparatively less for *Plot B* and *C*, with a *Percentage Error* less than 2%. Whereas, the error in detection is high for *Plot A* with *Percentage Error* of 12.81%. Further, on studying Figure 5.2 for both the ear detection methods, the error for *Plot A* is in most cases higher. This poor performance for one plot alone could be attributed to the variety of the plant. On visual analysis of the point cloud from *Plot A*, it was noticed that the size of the ears were bigger than the ones in the other two plots. This leads to scenarios of over-segmentation for the older crops, which causes an increase in the number of ears detected (as observed in Figure 5.1). This is confirmed further on studying the trends of the percentage error graphs in Figure 5.2.

Table 5.2: Evaluation of the effect of wheat variety on Ear Detection. (tested on mid-developmental stage)

| **Plot ID** | Voxel-Based | | Mean-Shift | |
|---|---|---|---|---|
| | Absolute Error ($ears/m^2$) | Percentage Error (%) | Absolute Error ($ears/m^2$) | Percentage Error (%) |
| Plot A | 41 | 12.81 | 73.86 | 19.83 |
| Plot B | 2.9 | 1.04 | 7.1 | 2.56 |
| Plot C | 6.4 | 1.77 | 3.4 | 0.94 |

Thus, to make the ear detection methods independent of the variety and developmental stages, strategy for adaptive selection of segmentation parameters is suggested.

### 5.2.3 Effect of Point Density

The input parameters of the two ear detection methods developed in this study are dependent on the point density of the dataset (as explained in Subsections 4.2.1 and 4.2.2). Hence, to understand the effect of point density, the detection approaches were first implemented on the downsampled datasets using the default input parameters identified by parameter tuning (as listed in Tables 4.2 and 4.1). As expected, their performance degraded with the decrease in point density. An example of this trend can be seen in Figure 5.3, which presents the number of ears detected for a subset of *Plot C* downsampled by retaining 75%, 50% and 25% of the original points. However, an interesting observation from the figure is that for the dataset downsampled to 75% of the original points, the performance of both the methods are consistent. Hence, this shows that both the ear detection methods are capable of handling a variation of upto 25% of the point density without adverse drop in their performance.

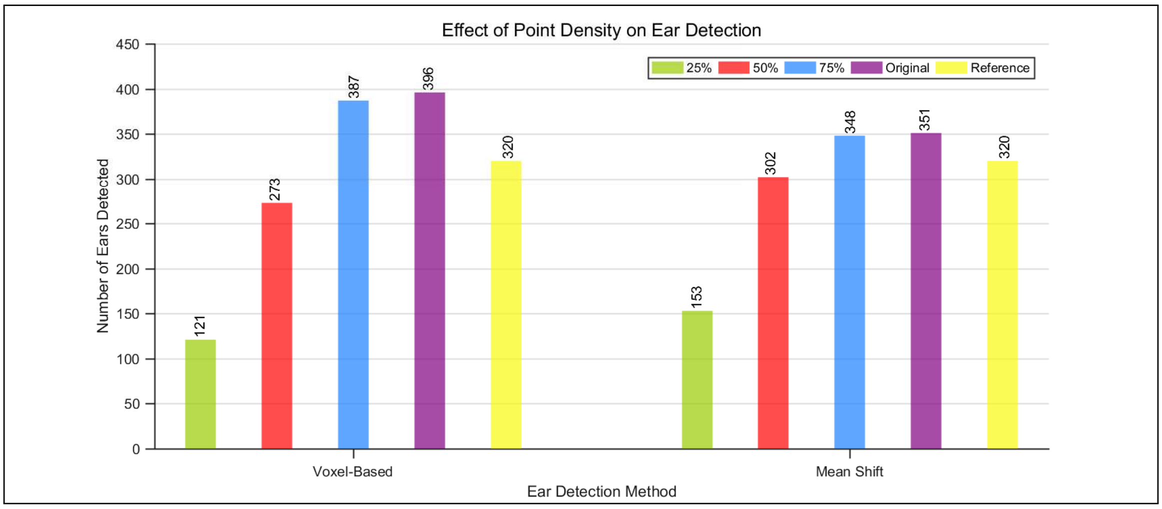



Figure 5.3: An example of the effect of point density in ear detection carried out using the default parameter values. Both the detection methods perform well when downsampled upto 75% of the original point density (A subset of *Plot C*, crops of age 225 days).

Next, another set of experiment was carried out where input parameters were chosen to according to the downsampled point density. Hence, for each of the nine downsampled subsets of the three plots, input parameter values as explained in Subsection 4.2.3 were used in the implementation.

The errors in ear detection for this experiment adapted to point density is shown in Table 5.3. It is seen from the table that the influence of point density on the voxel-based method is higher than that for the mean shift-based approach, especially for the $50\%$ and $25\%$ downsampled versions. This is because two of the four user-defined parameters are dependent on the point density of the dataset. But still the method gave reliable segmentation results when 50% of the original points were retained, even though there were cases of missing segments due to the thinning threshold. Thus, it is suggested to adapt the input parameter values according to the point density of the datasets being used. For the voxel-based approach applied on the dataset downsampled to 25% of the original point density, only a single iteration of point cloud thinning was carried out to calculate the errors shown in Table 5.3. Implementing a second iteration resulted in under-segmentation of the ears into small segments with less than 10 points and subsequent removal of these segments. Thus, it is recommended to use only one iteration of point cloud thinning when the dataset has a very low point density of $4points/cm^2$, i.e. 25% of the original point density used in this study.

For the mean shift-based ear detection, it is inferred from Table 5.3 that the influence of point density is comparatively lesser. This is because the clustering in mean-shift works based on the distance between the points, rather than on the point density. But still, the threshold used to remove small segments influences the results.

Table 5.3: Evaluation of the effect of Point Density on Ear Detection by downsampling the point cloud.

| **Percentage of Points Retained** | Voxel-Based | | Mean-Shift | |
|---|---|---|---|---|
| | RMSE $(ears/m^2)$ | MAPE(%) | RMSE $(ears/m^2)$ | MAPE (%) |
| 25 % | 87.57 | 24.29 | 62.78 | 15.89 |
| 50 % | 56.60 | 16.27 | 37.55 | 10.87 |
| 75 % | 52.07 | 14.7 | 58.66 | 16.50 |
| Original | 44.99 | 10.38 | 46.80 | 10.11 |

Thus, it is evident that the results from both the approaches developed are influenced to an extent by the point density of the dataset. However, from Figure 5.3, it is seen that the input parameters identified during the optimisation step are reliable when downsampled upto 75% of the original point density. In order to get reliable segmentation and classification results, it is recommended to select the input parameters based on the point density of the dataset. The optimal parameters selected by the cross validation method explained in Section 4.2 is is not universal and is limited to the data acquisition set-ups similar to the ones used in this study. However, based on the explanation for each parameter, the values can be modified for other experimental set-ups.

## 5.3 EVALUATION SUMMARY

This chapter evaluated the performance of the ear detection approaches in terms of different criteria. We observe that the performance of the ear detection methods is indeed influenced by the crop developmental stage and point density. Meanwhile, the variety of the crop is not as influential as the other two. Each of these criteria were found to influence the two ear detection methods in a different manner. It was found that the voxel-based method performs well for late developmental stages and should be implemented with adapted parameters for less dense point clouds. The analysis showed that the mean shift method was not as sensitive to the point density but leads to under-segmentation in overlapping areas.

## Chapter 6

# Results and Discussion

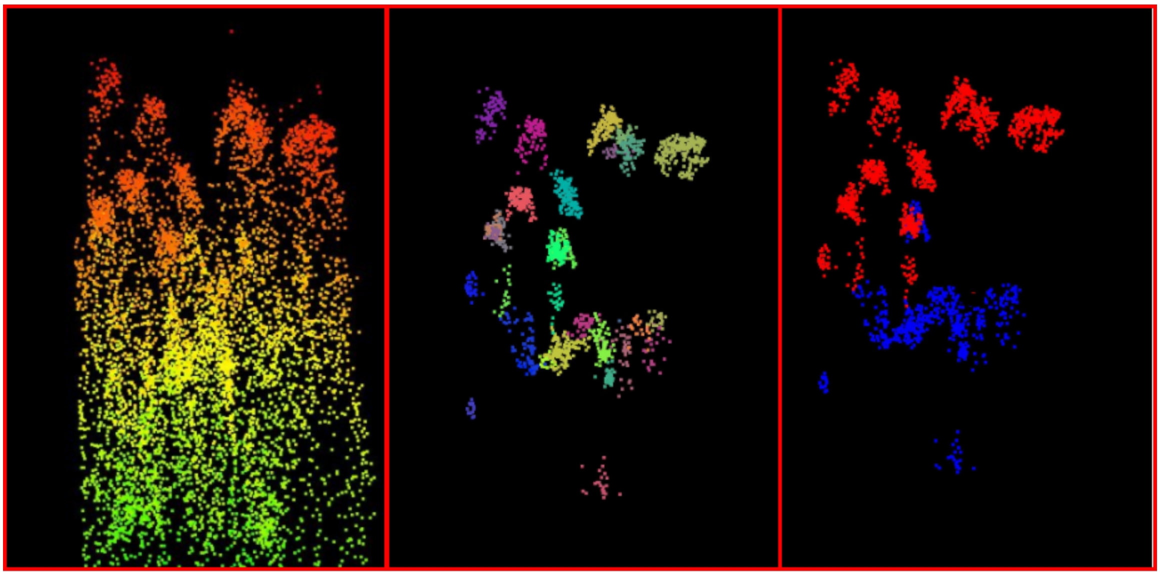
The results from the two ear detection methodology are presented in this chapter. We first start with a brief qualitative discussion of the segmentation results from the automated segmentation techniques used in ear detection. This is followed by an explanation of how the reference dataset was prepared by the manual segmentation of point clouds. Then results from the voxel-based and mean shift segmentation techniques are compared with each other and also with the manually labelled segments.

### 6.1 EAR DETECTION RESULTS

#### 6.1.1 Automatic Ear Detection Results

This section deals with a qualitative analysis of the ear segments extracted from the two automatic detection techniques. For each method, the raw point cloud, the segmentation results and the ear classification results are portrayed for a small sample subset.

##### Voxel-Based Ear Detection

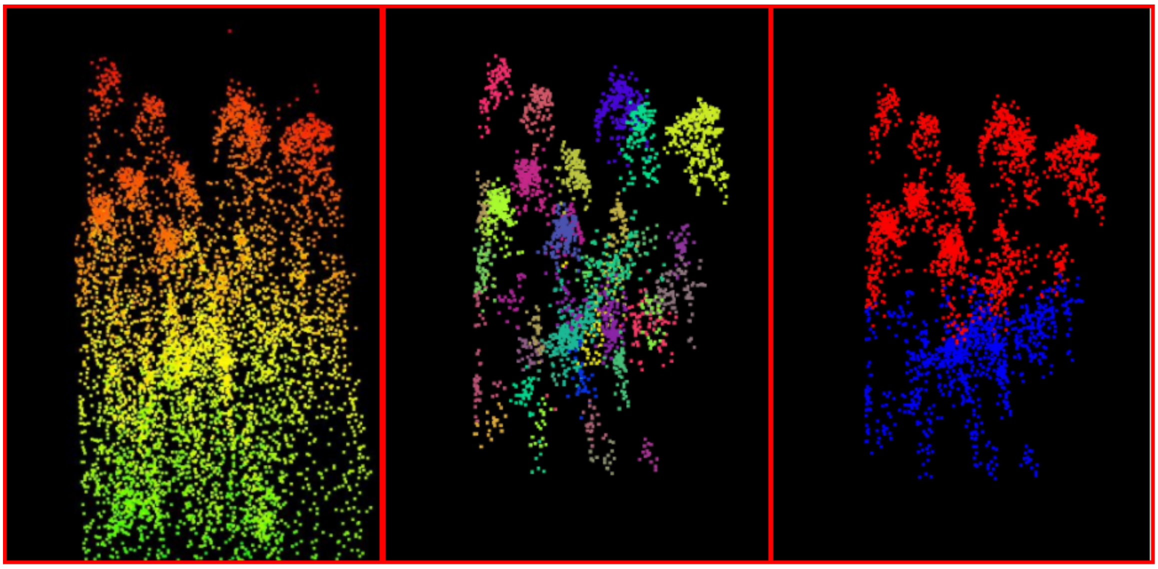


(a) (b) (c)

Figure 6.1: Voxel-Based Ear Detection results. The size of the ears is reduced due to point cloud thinning (a) Raw point cloud, colour coded based on height value (b) Voxel-based segmentation results with each segment displayed using a random colour (c) Ear classification where the red segments are the extracted ears (For a subset of *Plot A*, crops of age 225 days).

The thinning of point clouds in the voxel-based segmentation results in segments with very less number of points, when compared to the mean-shift segmentation. From Figure 6.1 (b), it can be seen that most of the noisy points surrounding the ears in the original point cloud have been removed. This results in the removal of few ear points as well. This ensures that the ear segments are distinctly identified even in regions of overlap. The points in directly connected voxels have been assigned the same segment number and each segment is displayed in a randomly assigned colour. In Figure 6.1 (c), we can see the finally extracted ear segments in red, while the non-ear segments are displayed in blue. It is noticed that the voxel-based method does not retain the size of the original ear and hence if used to estimate the size of the wheat ear, it will lead to underestimation. Thus, the thinning threshold should be reconsidered if the size of the wheat ear is needed for the study. However, for plant population estimation the size of the ear segments is not a concern.

#### Mean Shift-Based Ear Detection



(a) (b) (c)

Figure 6.2: Mean shift-Based Ear Detection results. The size of the ears is retained (a) Raw point cloud, colour coded based on height value (b) Mean shift-based segmentation results with each segment displayed with a random colour (c) Ear classification where the red segments are the extracted ears(For a subset of *Plot A*, crops of age 225 days).

The segmentation results from mean shift segmentation are representative of the raw point cloud i.e. the segments are not thinned and the size of the segments is retained. From Figure 6.2 (b), we notice that the segments retain majority of the points from the raw point cloud and hence are mostly noisy. Thus, the mean-shift based ear detection method can be directly extended to applications such as ear size estimation, without severe underestimation.

### 6.1.2 Manual Segmentation

This section explains how manual extraction of the reference segments was done and presents the comparison of ear detection results with manually prepared reference segments.

Table 6.1: Count of ears extracted by manual segmentation.

| **Plot** | **Subset Area** ($m^2$) | **Crop Age** | | |
|---|---|---|---|---|
| | | 194 days | 209 days | 225 days |
| A | 0.07 | 20 | 21 | 22 |
| B | 0.2 | 30 | 44 | 50 |
| C | 0.07 | 25 | 30 | 38 |

From the point clouds acquired on the three crop developmental stages, i.e. over crops of age 194, 209 and 225 days, a subset corresponding to the same area was taken from the *Plots A*, *B* and *C*. Since manual segmentation is a tedious task, very small subsets of area 7 to 20 $cm$ were chosen. From these 9 subsets, ear segments were manually segmented and labelled by 6 different operators using CloudCompare. The number of ears extracted by the operators for each subset is shown in Table 6.1. In order to check the variability between the operators, one common dataset ($Plot216$, crops of age 225 days) was provided to them.

From Table 6.1, we observe that the number of ears identified increases with the developmental stage of the plant. This concurs with our conclusion from Subsection 5.2.1 that with younger ear, the number of points is less and hence identification of ears becomes difficult.

### 6.1.3 Manual Segmentation Variability

An experiment was carried out to check the variability in manual segmentation among the different operators. All the six operators were given the same subset from *Plot A* acquired on crops 225 days old, to manually extract ear segments. Table 6.2 shows the number of ear segments extracted by each operator from the same subset. The varying results show that it is difficult even for a human operator to distinctly identify the ear segments from point cloud.

Table 6.2: Variability among different operators in manually extracting the ear segments(For a subset of *Plot C*, 225 days old). The subjectiveness of manual segmentation and difficulty to distinctly identify ear segments for humans is observed.

| Operator ID | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| Count | 27 | 19 | 17 | 19 | 14 | 22 |

### 6.1.4 Comparison with Manual Labels

The fitness of the two ear-detection methods was evaluated by comparing the number of ears observed by the automated methods with the reference dataset prepared by manual extraction. The results presented in this section were prepared using the optimal input parameter values identified in Table 4.2.

In addition to the RMSE and MAPE metrics described in Section 5.1, we used two more metrics, as explained below, that have been used in literature ( Yao et al., 2014; Y. Wang et al., 2016) to evaluate the performance of object detection algorithms.

$$Correctness = \frac{\text{Number of ears correctly detected}}{\text{Total number of ears detected}} \tag{6.1}$$

$$Completeness = \frac{\text{Number of ears correctly detected}}{\text{Number of reference ears}} \tag{6.2}$$

The two ear detection approaches were applied on the subsets on which the manual segmentation was performed. To calculate the detection error, manual counts from Table 6.1 were used as true values.

Table 6.3: Error in Ear Counts from the Automatic Detection methods calculated (Aggregated over the different plots and developmental stages).

| **Error Metric** | Voxel-Based | Mean-Shift |
|---|---|---|
| *RMSE* | 9.27 | 13.96 |
| *MAPE (%)* | 14.90 | 41.00 |
| *Correctness (%)* | 80.15 | 60.57 |
| *Completeness (%)* | 90.27 | 83.33 |

The *MAPE* values in Table 6.3 were calculated by comparing the number of ear segments identified by the automated ear detection methods with the number of ear segments manually extracted. It is observed that the voxel-based ear detection method has a lower *MAPE* of 14.90% whereas the mean shift-based ear detection has a value of 41.00%, after aggregation over different plots and developmental stages. It should be noted that before aggregation over different developmental stages, the voxel-based method had an average *MAPE* of 6% for crops aged 209 days and above. Whereas, the mean shift based method had an average *MAPE* of a very low 5% for crops of age 209 days and below.

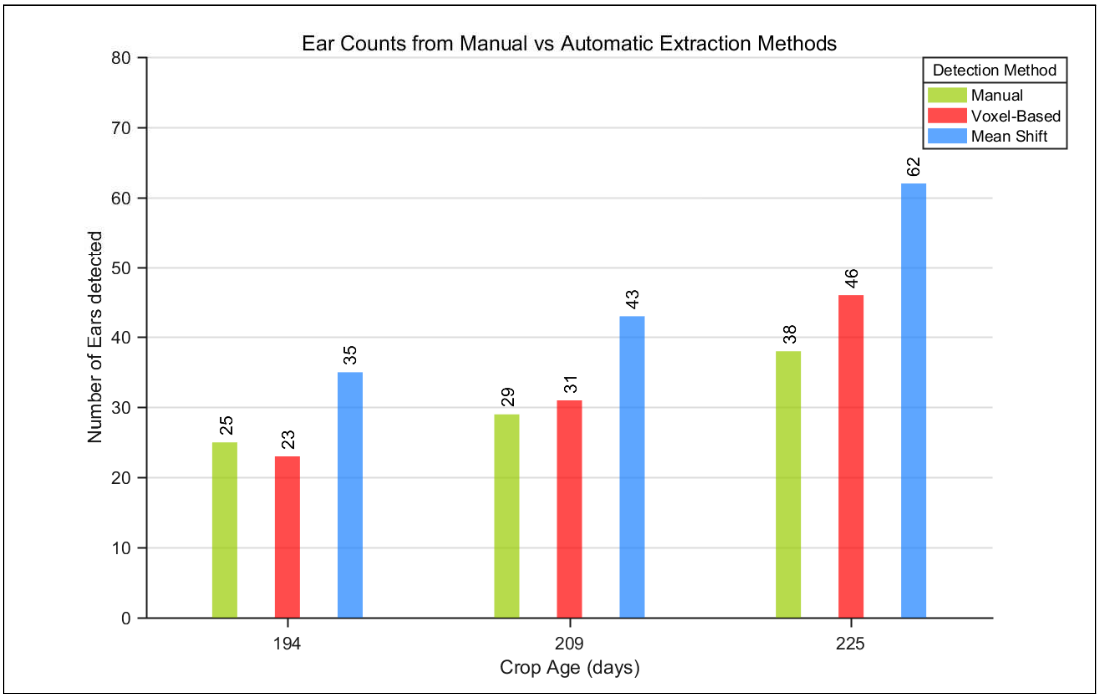



Figure 6.3: Comparing the number of ear segments detected from the Voxel-Based and Mean Shift-Based Ear Detection methods with manually created reference labels (A subset of *Plot C*)

This detection rate is an improvement when compared to the $0.9$ correlation with $40$ training points reported for image-based ear detection (Cointault et al., 2008). In a different study, (Saeys et al., 2009) reported a $94\%$ ear detection rate while using LiDAR observations with optimal acquisition settings. However, this was demonstrated on artificially constructed canopy and the proposed method was computationally very expensive.

The voxel-based detection method had a comparatively higher $Correctness$ value of $80.15\%$ while compared to the $60.57\%$ of the mean shift method. A similar trend is noticed for the $Completeness$ value for which the voxel-based method has a $90.27\%$ while the mean shift has only $83.33\%$. This indicates that the voxel-based approach is more successful than mean shift in identifying ear segments present in the scene with a relatively lower number of false ear segments.

Figure 6.3 presents the results from the two ear-detection methods for different developmental stages of the wheat crops in *Plot C*. From the graph, one can notice that the number of ears detected increases with the developmental stage of the crop. This holds for the manual segmentation as well, which means more ears were identified by the operator in the late flowering stage when compared to the early stage. This concurs with the observations made in section 5.2.1 regarding the change in trends of ear detection with the developmental stage.

## 6.2 QUALITATIVE OBSERVATIONS

### 6.2.1 Segmentation Quality

Besides the quantitative comparisons, a visual comparison was also carried out to check the quality of the extracted ear segments. Figure 6.4 shows a comparison between the manually extracted ear segments and the segments from the voxel based and mean shift segmentation methods. One striking difference in the segments obtained from the two methods is the size of the segments. For the voxel-based segmentation, the segments are usually smaller due to the thinning of the point cloud. Whereas, for the mean shift segmentation, the size and shape of the ears is retained; if there were some noisy points close to the object, they are retained as well. This can be seen in Figure 6.4 (a).

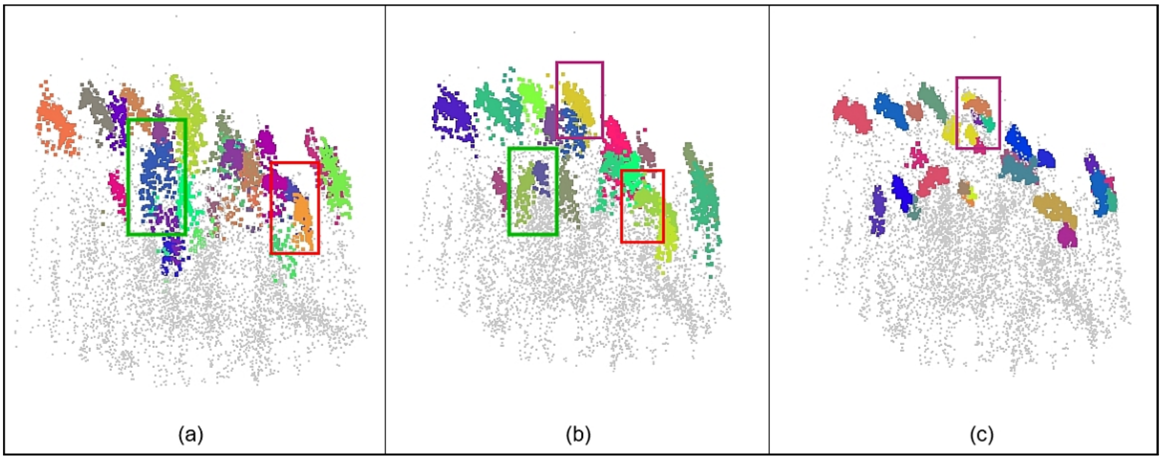



Figure 6.4: Quality of Segments obtained from (a)Mean-Shift Segmentation (b) Manual Segmentation (c)Voxel-Based Segmentation(For a subset of *Plot A*, crops aged 225 days)

The red boxes in 6.4 (a) and (b) highlight an example of over-segmentation of the point cloud by mean shift segmentation. This concurs with the claims made in Subsection 5.2.1 regarding the over-segmentation of ear segments by mean-shift segmentation for late developmental stages.

The green boxes in 6.4 (a) and (b) bring to attention a case of under-segmentation by mean shift segmentation in overlapping areas. This is because the ghost returns present close to the ear and leaf objects are retained during the filtering applied before the mean shift technique; as a result the areas of overlap get clustered as single segments. This could be overcome by using adaptive kernel bandwidths selected based on point density.

The purple boxes in Figure 6.4 (b) and (c) highlight an example of over-segmentation by the voxel-based detection method. This is observed for wheat ears with stunted growth that are small in size. One possible way to avoid this would be to carry out a single iteration of voxelization with automated voxel size selection according to the local point density of the point cloud.

#### Mean Shift-Based Ear Detection:

### 6.2.2 Processing Time

In terms of computational time, the voxel-based approach is faster between the two. The entire implementation of voxel-based approach was developed and carried out in Matlab and takes approximately 20 minutes to segment and classify one micro-plot of roughly 3.2 million points. On the other hand, the mean shift on components approach implemented using a combination of Matlab and C++, takes more or less 2 hours to run on the full micro-plot. It should be noted that this is an almost 80% reduction in computational time when compared to running the mean shift directly on the point cloud. However, in terms of processing time, the voxel-based approach is preferable and if needed, could be optimized to be implemented on-line on the scanner system since it is now able to process 50,000 points in less than 30 seconds. The comparisons regarding computational time have been made by running the processes on the same computer.

## 6.3 DISCUSSION

### 6.3.1 Usability of Intensity Information

The laser scanner used in this study records not only the 3D positional data, but also the intensity value that depends on the reflectivity of the surface. We decided to test the suitability of this intensity data for distinguishing between ear and non-ear segments.

#### Intensity Correction

The intensity values recorded by the sensor are affected by several factors such as the influence of atmosphere, instrumental effects, acquisition geometry and target characteristics(Kashani et al., 2015). Hence, a simple range-based intensity correction normalised to a user-defined standard range, demonstrated for diffuse targets such as in forestry applications (Donoghue et al., 2007; García et al., 2010) was applied using the following equation:

$$I' = I * \frac{R^2}{{R_s}^2} \tag{6.3}$$

$I'$ = Normalised intensity.
$I$ = Raw intensity value.
$R$ = Range (distance between the sensor and the object surface).
$R_s$ = User-defined standard range for intensity normalisation.

After normalization, the ground points were removed and the histogram of the normalized intensity for a ear-classified dataset (*Plot A* acquired on crops aged 225 days) as shown in Figure 6.5

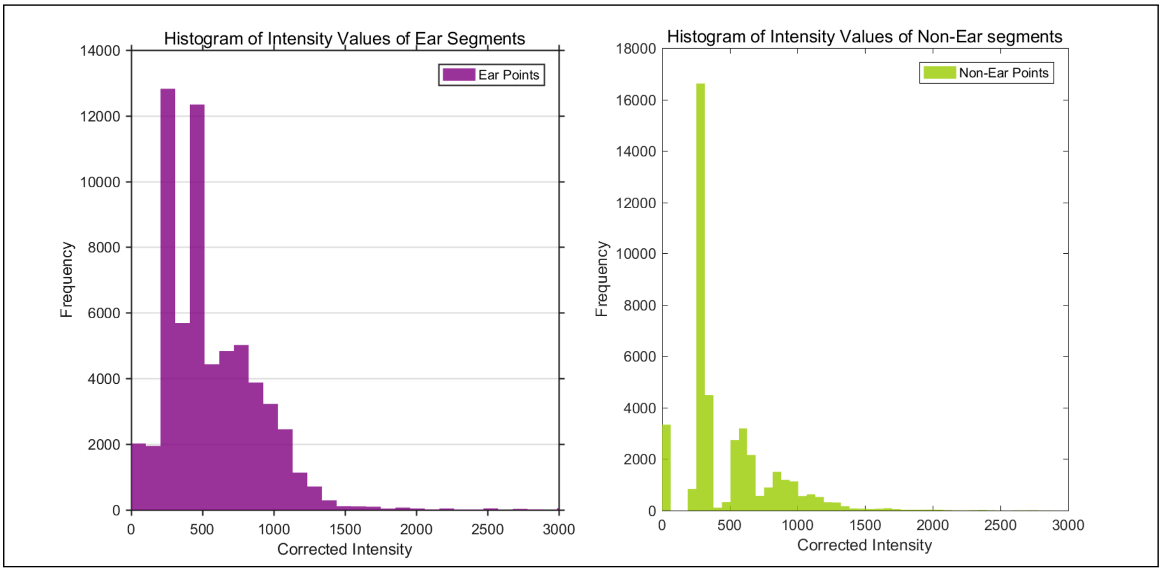
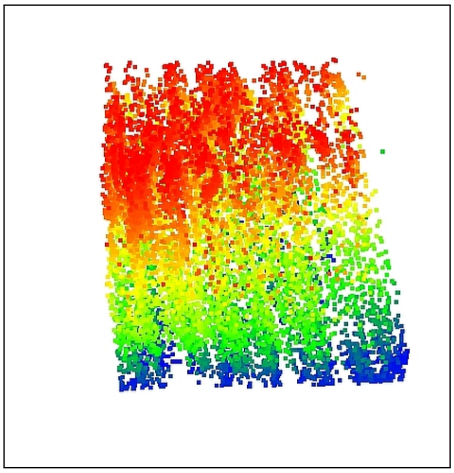



Figure 6.5: Histogram of the corrected intensity values for the ear and non-ear points. The range of intensity values is almost the same which implies these values cannot be used for distinguishing ear points from non-ear points. No distinct peaks are observed (A subset of *Plot A*, acquired over 225 days old crops).

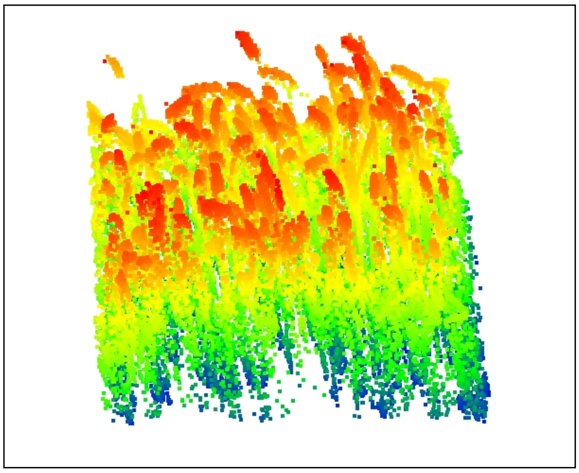
was analysed. The figure shows the histograms of the ear and non-ear points after the intensity correction. It is evident that there is no distinct value that could be used to distinguish between the ear and non-ear segments as the range of the intensity values is almost the same. Different combinations of thresholds were tested to distinguish between the ear and non-ear points. But the thresholds always resulted in a combination of ear and non-ear points. This is because the intensity value is suitable to distinguish between plant and ground points (Hofle, 2014) but is not sufficient to distinguish between the different plant parts. Moreover, during the early developmental stage, majority of the leaves and the ears are green. Similarly, during the harvest stage, majority of the leaves, including the wheat ear, attain senescence and turn brown in colour. This makes the classification between different plant organs based on intensity difficult.

### 6.3.2 Influence of Lidar Look Angle

During data acquisition, laser scanners fitted at nadir, looking downward to observe the crop canopy and fitted laterally, with an inclination of 45°, were used (as described in Figure 3.1). The datasets from these two positions were qualitatively assessed. It was found that the point density of the lateral dataset was only 15% that of the nadir dataset. For fully grown crops, the lateral perspective of the sensor is able to observe only the plants directly in front of the sensor as the plants behind are overshadowed by the ones present in the first row. As a result, there are hardly any points that are returned by the points in the last row. The nadir dataset, on the other hand, provides a better description of all the crops under its field of view. Though the number of ghost return points increase with distance from the sensor, the datasets from the nadir perspective provide reliable representation of the crop canopy architecture. This can be observed in Figure 6.6 which shows the LiDAR datasets acquired simultaneously from the two different perspectives, colour coded based on their height values. The data in Figure 6.6 (a) appears like a haze of noisy points whereas the nadir dataset in Figure 6.6 (b) depicts the structure on top of the canopy in detail.



(a) Points acquired with lateral look angle of 45°. It is difficult to decipher the plant structure from this point cloud.



(b) Points acquired with downward nadir look angle. The structure of wheat ears is easy identifiable.

Figure 6.6: An example showing the difference between the two perspectives in representing the canopy architecture.

Thus, for the detection of wheat ears from fully grown crops in the field, lateral datasets do not provide sufficient description of the canopy. Hence, the nadir datasets are preferable.

## 6.4 SUMMARY

The segmentation quality was assessed using reference segments extracted manually. The voxel-based segmentation showed a higher *Completeness* and *Correctness* value of 90.27% and 80.15% respectively aggregated over the different developmental stages and varieties. Analysing the subjectiveness of the operators during manual segmentation showed that the extraction of ear segments is a tricky task for humans as well. An experiment conducted to check the suitability of intensity information to classify ear and non-ear segments revealed that it is not possible to classify ears segments based on intensity values.

Following are the important observations made during the qualitative analysis and comparison of the ear detection methods:

- The voxel-based approach performs well for crops of age 209 days and older where the wheat ears are large in size. For early developmental stage (194 days and younger), there are situations where most of the points on the wheat ears are filtered during the point cloud thinning procedure. The mean shift-based approach, on the other hand, is not dependent on the point density during the segmentation step. This leads to reliable segmentation results where the shape of the wheat ears are preserved.

- The voxel-based approach performs well even in areas of high overlap as the point cloud thinning leads to removal of points in less dense areas such as the stem. But this is not the case for the mean-shift which does not perform well in overlapping areas. This is because the filtering technique used in this approach removes only the hovering points.

- Since the voxel-based method of ear detection is sensitive to point density of the dataset, the input parameter values need to be adapted as per the plot characteristics.

- Similarly, using an adaptive kernel bandwidth based on point density for mean-shift segmentation will help in avoiding under-segmentation issues.
- The performance of both the methods is consistent up to $\pm 25\%$ difference in the point density. For larger differences in point density, the performance is degraded.

Hence, these evaluations helped to understand the two detection methods individually and the factors they are sensitive to. In the next chapter, the final conclusions and recommendations are made while addressing the research questions raised in Chapter 1 during the beginning of the study.

## Chapter 7

# Conclusion and Recommendations

## 7.1 CONCLUSION

In this study, we designed, analysed and evaluated two approaches for the detection of wheat ears from laser scanned 3D point clouds. The task of ear detection was split into two steps - segmentation and classification of segments. For segmentation of the point cloud, voxel-based segmentation and mean-shift segmentation methods were chosen and implemented. For classification of the segments, an ear classification methodology was developed taking into consideration the position of wheat ears in the canopy, overlap between plants and presence of stunted plants within the canopy. The performance of these two ear detection method was then assessed based on developmental stage of the crop, wheat variety and point density. It was noted that the two methods proposed have in most cases ear detection errors of roughly 15%. The voxel-based approach performs well for late developmental stages with a comparatively very short computational time. Meanwhile, the mean-shift method performs well for different point densities and gives reliable results for early developmental stages and has a higher processing time. Since both the ear-detection approaches are in some way influenced by the point density of the dataset, the input detection parameters need to be adapted to achieve good count results. Another interesting observation was the variation in the number of ears extracted by different operators by manual segmentation which focuses on the fact ear detection is a difficult task even for a human operator. Depending on the available prior knowledge regarding the crops studied and accuracy required for the application, one of the two ear detection approaches may be selected. Thus, through this study, two ear detection techniques that depend solely on the 3D coordinates of the points describing wheat ears have been developed and compared.

### 7.1.1 Research Questions: Answered

This section discusses the solutions obtained during the project execution for the research questions raised at the beginning of the research.

1. What are the methods that have already been proposed in literature for the segmentation of plant organs?

   The methods that already exist in literature for segmentation of plant organs is discussed in Chapter 2. The methods reviewed were demonstrated on individual plants under laboratory conditions, where the effects of overlapping plants and ghost returns are minimal. The methods proposed for the detection of individual plants in the field were also reviewed. In most cases, machine learning methods have been proposed to train classifiers and segment the plant parts using these classifiers. However, for this study it was preferred to develop an unsupervised method that can be adapted to different crop varieties and types. The voxel-based segmentation approach was chosen from among these as it gives the flexibility to perform a localised point cloud thinning and segmentation, based on point density, in an efficient and quick manner.

2. What are the methods in literature for single tree detection from airborne laser scanned data that are relevant to the proposed study?

   Chapter 2 reviews the methods demonstrated for the individual detection of trees from airborne laser scanner as well. Many of these methods propose a surface growing approach or watershed segmentation that use a local maxima as the starting point of clustering. Detecting local maxima for each plant in the field would be a difficult task, given the random cropping pattern of the wheat crop and ghost returns hovering over the canopy. Among these methods, the mean shift segmentation was chosen since this could be customised to the shape of the object being detected and does not require a priori information such as an initial starting point or number of classes in the dataset.

3. How can the knowledge on how wheat seeds are sown (the number of rows, average spacing between each plant and rows) and how they look like be incorporated in the segmentation algorithm?

   Generally, the wheat plants are cropped with random spacing between them. This average space between the plants is used to decide the size of the voxels used in the voxel-based segmentation. The size of the wheat ear increases with the age of the crop which naturally causes an increase in the number of points describing the wheat ear. Thus, the thinning thresholds for points removal may be decided based on the age of the crop to get good segmentation results. For the mean shift segmentation, the bandwidth parameters for clustering are fixed depending on the average size of the wheat ear. Once again, the age of the crop may be used to decide the input parameters for segmentation. Though the orientation of rows in the field can be identified, the plants are not organized in a regular manner in these rows. Hence, information regarding the orientation and number of rows may not add additional value to the segmentation approaches adopted in this study. A detailed explanation on how to decide the input parameters for the segmentation process is described in Subsections 4.2.1 and 4.2.2 for the voxel-based and mean shift ear detection approaches respectively.

4. What is the size and distribution of the noise and how can it be reduced?

   The different types of noise present in the dataset is discussed in Subsection 3.1.3. The noise present in the point cloud is distributed in a random manner and are due to the multi-edge effect. These were removed by a voxel dependent thinning for the voxel-based segmentation (described in Subsection 3.2.1) which removed the hovering points around the objects and also most of the points in the stem region. For the mean shift segmentation, the noisy points were removed by a threshold decided based on point density in a cylindrical neighbourhood for each point in the point cloud (described in Appendix A). For the dataset used in this study, approximately 5% of the total number of points were noisy points and were eliminated in the filtering steps before applying the segmentation techniques.

5. What is the influence of the shadow returns (resulting in noisy objects) on the segmentation process?

   Executing the segmentation on the point cloud, without prior filtering resulted in poor segments. Since the voxel-based segmentation looks for connected components in the point cloud, there were several small components and several large components connected due to the ghost return points surrounding the objects. Though the mean shift segmentation was able to handle the noisy points, running the segmentation without removing the ghost returns almost doubled the computational time. Most of the resulting segments were correctly segmented but some cases of under-segmentation was noticed.

6. How does the performance of the segmentation algorithm differ depending on the angle of acquisition (nadir and lateral)? Which dataset (nadir or lateral) is preferable to estimate the count of ears?

   For the detection of ear organs from fully grown wheat plants, it is suitable to use the nadir dataset. When a lateral angle is used for acquisition, only the objects closest to the sensor is observed clearly; the objects behind are overshadowed by the ones in the front. Also, in our study, the optimal operating range for the laser scanner used was 0.7 $m$ to $3m$. Hence, for the lateral dataset, the range between the sensor and the plants in the last row of the plot was high resulting in scattered points. As the architecture of the plant is poorly recorded, attempts on segmenting the point cloud is futile. This is discussed in Subsection 6.3.2.

7. How does the point density affect the segmentation performance?

   The effect of point density on the segmentation performance was evaluated and discussed in detail in Subsection 5.2.3. Both the ear detection approaches proposed are dependent on the point density of the datasets. In order to get optimal results, it is recommended to select input parameters that would suit the point density of the dataset. This means, for a point cloud with point density less than $16 points/cm^2$ (the point density of the datasets used in this study), thresholds that are lenient compared to the ones detected for this study are recommended for the removal of points and small segments. The Table 4.3 lists how the optimal parameters were modified to suit the downsampled datasets.

8. Will the addition of intensity information from the laser scanner complement the counting accuracy?

   From a simple experiment, it was concluded that the addition of intensity information from the laser scanner will not be helpful to improve the classification between ear and non-ear segments. This is explained using the corrected intensity histogram in Figure 6.5.

9. What is the effect of the variety and senescence of wheat ears on the segmentation performance?

   Sections 5.2.1 and 5.2.2 give a detailed account of how the temporal variability and variety of the wheat crops affect the segmentation and in turn the ear detection performance. In general, there is an underestimation in the number of ear segments for the early developmental stage for the voxel-based approach. This trend was noticed in the manual segmentation results as well which could be attributed to the comparatively lesser number of points on the ear segments. The mean shift segmentation results in under-segmentation in overlapping areas of late developmental stages. A suggestion to overcome these drawbacks would be to use adaptive input parameter values selected depending on the point density.

## 7.2 RECOMMENDATIONS

The following suggestions are recommended for future research work:

- Investigate the use of adaptive kernel bandwidths for mean shift segmentation decided based on weights calculated from the neighbouring point density and heights of the points. This helps to avoid under-segmentation of ears when there is overlap between them

- Attempts should be made to extend the filtering technique used in the mean shift segmentation to remove points on the stem region. This will avoid the misclassification for early developmental stages explained in Subsection 5.2.1.

- Inspections should be made to make the ear classification procedure independent of the height thresholding step (explained in Subsection 3.2.3). This will help in detecting stunted ears that are present below the average ear layer in the canopy. One way to execute this would be to test the suitability of colour information from RGB images acquired by the cameras mounted with the laser scanner (as shown in Figure 3.1) to improve the ear classification accuracy.

# References

Bucksch, A., Lindenbergh, R., Menenti, M., & Rahman, M. Z. (2009). Skeleton-based botanic tree diameter estimation from dense LiDAR data. In U. N. Singh (Ed.), *Proceedings of spie - the international society for optical engineering* (Vol. 7460, p. 746007). doi: 10.1117/12.825997

Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? - Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, *7*(3), 1247–1250. doi: 10.5194/gmd-7-1247-2014

Cointault, F., Guerin, D., Guillemin, J., & Chopinet, B. (2008). In-field triticum aestivum ear counting using colour texture image analysis. *New Zealand Journal of Crop and Horticultural Science*, *36*(2), 117–130. doi: 10.1080/01140670809510227

Comaniciu, D., & Meer, P. (2002). Mean shift: A robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *24*(5), 603–619. doi: 10.1109/34.1000236

Deery, D., Jimenez-Berni, J., Jones, H., Sirault, X., & Furbank, R. (2014). Proximal Remote Sensing Buggies and Potential Applications for Field-Based Phenotyping. *Agronomy*, *4*(3), 349–379. doi: 10.3390/agronomy4030349

Donoghue, D., Watt, P., Cox, N., & Wilson, J. (2007). Remote sensing of species mixtures in conifer plantations using LiDAR height and intensity data. *Remote Sensing of Environment*, *110*(4), 509–522. doi: 10.1016/j.rse.2007.02.032

Eitel, J. U., Magney, T. S., Vierling, L. A., Brown, T. T., & Huggins, D. R. (2014). LiDAR based biomass and crop nitrogen estimates for rapid, non-destructive assessment of wheat nitrogen status. *Field Crops Research*, *159*, 21–32. doi: 10.1016/j.fcr.2014.01.008

FAO. (March 2016). *GIEWS - Crop Prospects and Food Situation, No.1.* Retrieved June 1, 2016, from `http://www.fao.org/giews/english/cpfs/I5455e/I5455E.html`

Ferraz, A., Bretar, F., Jacquemoud, S., Goncalves, G., & Pereira, L. (2010). 3D segmentation of forest structure using a mean-shift based algorithm. In *2010 IEEE International Conference on Image Processing* (pp. 1413–1416). IEEE. doi: 10.1109/ICIP.2010.5651310

Fukunaga, K., & Hostetler, L. (1975). The estimation of the gradient of a density function, with applications in pattern recognition. *IEEE Transactions on Information Theory*, *21*(1), 32–40. doi: 10.1109/TIT.1975.1055330

Furbank, R. T., & Tester, M. (2011). Phenomics–technologies to relieve the phenotyping bottleneck. *Trends in Plant Science*, *16*(12), 635–44. doi: 10.1016/j.tplants.2011.09.005

García, M., Riaño, D., Chuvieco, E., & Danson, F. M. (2010). Estimating biomass carbon stocks for a Mediterranean forest in central Spain using LiDAR height and intensity data. *Remote Sensing of Environment*, *114*(4), 816–830. doi: 10.1016/j.rse.2009.11.021

Garrido, M., Paraforos, D., Reiser, D., Vázquez Arellano, M., Griepentrog, H., & Valero, C. (2015). 3D Maize Plant Reconstruction Based on Georeferenced Overlapping LiDAR Point Clouds. *Remote Sensing*, *7*(12), 17077–17096. doi: 10.3390/rs71215870

Gorte, B., Oude Elberink, S., Sirmacek, B., & Wang, J. (2015, aug). IQPC 2015 TRACK: TREE SEPARATION AND CLASSIFICATION IN MOBILE MAPPING LIDAR DATA. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XL-3/W3*, 607–612. doi: 10.5194/isprsarchives-XL-3-W3-607-2015

Hofle, B. (2014). Radiometric Correction of Terrestrial LiDAR Point Cloud Data for Individual Maize Plant Detection. *IEEE Geoscience and Remote Sensing Letters*, *11*(1), 94–98. doi: 10.1109/LGRS.2013.2247022

Hosoi, F., & Omasa, K. (2006). Voxel-Based 3-D Modeling of Individual Trees for Estimating Leaf Area Density Using High-Resolution Portable Scanning Lidar. *IEEE Transactions on Geoscience and Remote Sensing*, *44*(12), 3610–3618. doi: 10.1109/TGRS.2006.881743

Hosoi, F., & Omasa, K. (2009). Estimating vertical plant area density profile and growth parameters of a wheat canopy at different growth stages using three-dimensional portable lidar imaging. *ISPRS Journal of Photogrammetry and Remote Sensing*, *64*(2), 151–158. doi: 10.1016/j.isprsjprs.2008.09.003

Houldcroft, C., Campbell, C., Davenport, I., Gurney, R., & Holden, N. (2005). Measurement of canopy geometry characteristics using LiDAR laser altimetry: a feasibility study. *IEEE Transactions on Geoscience and Remote Sensing*, *43*(10), 2270–2282. doi: 10.1109/TGRS.2005.856639

Journaux, L., Marin, A., Cointault, F., & Miteran, J. (2010). Fourier Filtering for Wheat Detection in a Context of Yield Prediction. In *XVIIth World Congress of the International Commission of Agricultural and Biosystems Engineering (CIGR).* Québec City, Canada.

Kashani, A., Olsen, M., Parrish, C., & Wilson, N. (2015). A Review of LIDAR Radiometric Processing: From Ad Hoc Intensity Correction to Rigorous Radiometric Calibration. *Sensors*, *15*(11), 28099–28128. doi: 10.3390/s151128099

Knapton, S. (2016). *Genetically modified wheat could be grown in Britain from next spring.* Retrieved 2016-11-07, from `http://www.telegraph.co.uk/science/2016/11/04/genetically-modified-wheat-could-be-grown-in-britain-from-next-s/`

Koenig, K., Höfle, B., Hämmerle, M., Jarmer, T., Siegmann, B., & Lilienthal, H. (2015). Comparative classification analysis of post-harvest growth detection from terrestrial LiDAR point clouds in precision agriculture. *ISPRS Journal of Photogrammetry and Remote Sensing*, *104*, 112–125. doi: 10.1016/j.isprsjprs.2015.03.003

LemnaTec. (2015, November). *Digital Field Phenotyping* (Tech. Rep.). Author. doi: 10.13140/RG.2.1.1949.6404

Li, W., Guo, Q., Jakubowski, M. K., & Kelly, M. (2012). A New Method for Segmenting Individual Trees from the Lidar Point Cloud. *Photogrammetric Engineering & Remote Sensing*, *78*(1), 75–84. doi: 10.14358/PERS.78.1.75

Lin, Y. (2015). LiDAR: An important tool for next-generation phenotyping technology of high potential for plant phenomics? *Computers and Electronics in Agriculture*, *119*, 61–73. doi: 10.1016/j.compag.2015.10.011

Lumme, J., Karjalainen, M., Kaartinen, H., Kukko, A., Hyyppä, J., Hyyppä, H., … Kleemola, J. (2008). Terrestrial laser scanning of agricultural crops. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *37*, 563–566.

Melzer, T. (2007). Non-parametric segmentation of ALS point clouds using mean shift. *Journal of Applied Geodesy*, *1*(3). doi: 10.1515/jag.2007.018

Miller, T. D. (1992). Texas Agricultural Extension Service Growth Stages of Wheat: Identification and Understanding Improve Crop Management. *Better Crops With Plant Food*, *76*(3), 12–17.

Nguyen, A., & Le, B. (2013). 3D point cloud segmentation: A survey. In *2013 6th IEEE Conference on Robotics, Automation and Mechatronics (RAM)* (pp. 225–230). IEEE. doi: 10.1109/RAM.2013.6758588

Otsu, N. (1979). A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, *9*(1), 62–66. doi: 10.1109/TSMC.1979.4310076

Paproki, A., Fripp, J., Salvado, O., Sirault, X., Berry, S., & Furbank, R. (2011). Automated 3D

Segmentation and Analysis of Cotton Plants. In *2011 International Conference on Digital Image Computing: Techniques and Applications* (pp. 555–560). Noosa, QLD: IEEE. doi: 10.1109/DICTA.2011.99

Paulus, S., Dupuis, J., Mahlein, A.-K., & Kuhlmann, H. (2013). Surface feature based classification of plant organs from 3D laserscanned point clouds for plant phenotyping. *BMC bioinformatics*, *14*, 238. doi: 10.1186/1471-2105-14-238

Paulus, S., Dupuis, J., Riedel, S., & Kuhlmann, H. (2014). Automated Analysis of Barley Organs Using 3D Laser Scanning: An Approach for High Throughput Phenotyping. *Sensors*, *14*(7), 12670–12686. doi: 10.3390/s140712670

Ramamurthy, B., Doonan, J., Zhou, J., Han, J., & Liu, Y. (2015). Skeletonization of 3D Plant Point Cloud Using a Voxel Based Thinning Algorithm. *European Signal Processing Conference (EUSIPCO)*(23), 2686–2690. Retrieved from `http://www.eusipco2015.org/`

Saeys, W., Lenaerts, B., Craessaerts, G., & De Baerdemaeker, J. (2009). Estimation of the crop density of small grains using LiDAR sensors. *Biosystems Engineering*, *102*(1), 22–30. doi: 10.1016/j.biosystemseng.2008.10.003

Scott, A. (2014). *Global food security: could wheat feed the world? | Global Development Professionals Network | The Guardian.* Retrieved 2016-06-01, from `http://www.theguardian.com/global-development-professionals-network/2014/apr/01/international-wheat-yield-partnership-food-security`

SICK Germany. (2007). *The LMS400 Laser Measurement System* (Tech. Rep.).

Sinclair, T., & Jamieson, P. (2006). Grain number, wheat yield, and bottling beer: An analysis. *Field Crops Research*, *98*(1), 60–67. doi: 10.1016/j.fcr.2005.12.006

Sionit, N., Hellmers, H., & Strain, B. R. (1980). Growth and Yield of Wheat under CO2 Enrichment and Water Stress1. *Crop Science*, *20*(6), 687. doi: 10.2135/cropsci1980.0011183X002000060003x

Sirmacek, B., & Lindenbergh, R. (2015). Automatic Classification of Trees from Laser Scanning Point Clouds. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, *II-3/W5*, 137–144. doi: 10.5194/isprsannals-II-3-W5-137-2015

Solberg, S., Naesset, E., & Bollandsas, O. M. (2006). Single Tree Segmentation Using Airborne Laser Scanner Data in a Structurally Heterogeneous Spruce Forest. *Photogrammetric Engineering & Remote Sensing*, *72*(12), 1369–1378. doi: 10.14358/PERS.72.12.1369

UN DESA. (2015). *World population projected to reach 9.7 billion by 2050.* Retrieved 2016-08-26, from `http://www.un.org/en/development/desa/news/population/2015-report.html`

Van Genchten, B., Caner, H., Heine, E., García, J. L. L., Poelman, R., & Quintero, M. S. (2008). *Theory and practice on Terrestrial Laser Scanning*. Universidad Politecnica de Valencia Editorial.

Vauhkonen, J., Ene, L., Gupta, S., Heinzel, J., Holmgren, J., Pitkanen, J., … Maltamo, M. (2012). Comparative testing of single-tree detection algorithms under different types of forest. *Forestry*, *85*(1), 27–40. doi: 10.1093/forestry/cpr051

Vosselman, G., Gorte, B., Sithole, G., & Rabbani, T. (2004). Recognising Structure in Laser Scanner Point Clouds. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, *46*(April 2016), 33–38. doi: 10.1002/bip.360320508

Wahabzada, M., Paulus, S., Kersting, K., & Mahlein, A.-K. (2015). Automated interpretation of 3D laserscanned point clouds for plant organ segmentation. *BMC bioinformatics*, *16*(1), 248. doi: 10.1186/s12859-015-0665-2

Wang, J., & Shan, J. (2009). Segmentation of LiDAR point clouds for building extraction. *American Society for Photogrammetry and Remote Sensing*, 9–13.

Wang, Y., Cheng, L., Chen, Y., Wu, Y., & Li, M. (2016). Building Point Detection from Vehicle-Borne LiDAR Data Based on Voxel Group and Horizontal Hollow Analysis. *Remote Sensing*, *8*(5), 419. doi: 10.3390/rs8050419

Wang, Y., Weinacker, H., & Koch, B. (2008). A Lidar point cloud based procedure for vertical canopy structure analysis and 3D single tree modelling in forest. *Sensors*, *8*(6), 3938–3951. doi: 10.3390/s8063938

Weinmann, M., Mallet, C., & Brédif, M. (2016). Detection, Segmentation and Localization of Individual Trees from MMS Point Cloud Data. In *Geobia 2016.* Enschede.

Weiss, U., & Biber, P. (2011). Plant detection and mapping for agricultural robots using a 3D LIDAR sensor. *Robotics and Autonomous Systems*, *59*(5), 265–273. doi: 10.1016/j.robot.2011.02.011

Yao, W., Krull, J., Krzystek, P., & Heurich, M. (2014). Sensitivity analysis of 3D individual tree detection from LiDAR point clouds of temperate forests. *Forests*, *5*(6), 1122–1142. doi: 10.3390/f5061122

Yao, W., Krzystek, P., & Heurich, M. (2012). Tree species classification and estimation of stem volume and DBH based on single tree extraction by exploiting airborne full-waveform LiDAR data. *Remote Sensing of Environment*, *123*, 368–380. doi: 10.1016/j.rse.2012.03.027

## Appendix A

# Number of Points per Ear

The following sections describe the procedure that was used to estimate the average number of 3D points describing an ear organ in the dataset used in the study. This method was used in the mean-shift based ear detection method to decide thresholds for removing multi-edge returns from adjacent plant organs found in the point cloud. The trial and error experiments carried out to select a filtering threshold are also explained.

### A.1 METHOD DESCRIPTION

The dataset has hovering points in and around the plant canopy due to the averaging of the multi-edge returns from the plant structures. Removing these points helps to speed up the clustering during mean shift segmentation. In order to identify the non-ear points, the following steps were carried out.

- For each point in the point cloud, a cylindrical neighbourhood of height 6 cm and radius 3 cm was defined, with that point as centre. A cylinder was chosen since the ears roughly follow the shape of a cylinder. The dimensions of the cylinder were chosen on a trial and error basis to approximately follow the average size of a wheat ear.
- For each point, the number of neighbours within the 3D cylindrical neighbourhood was calculated and stored in a new attribute field.
- The point cloud was then visualised based on the values stored in the new attribute field, i.e. the number of neighbours for each point, as shown in Figure A.1. This helped to identify the average number of points per ear.

### A.2 THRESHOLD SELECTION

Based on a trial and error method, several range of values were tested to identify an approximate number of points per ear. When the points with values between 35 to 120 were grouped together, all the wheat ears with a mix of points in the stem region could be identified. This implies, that points with less than 35 neighbours were non-ear points, as shown in blue in the Figure A.1. This was logical since the hovering points do not have high point density in their local neighbourhood. Hence, it was decided to use a lower threshold of 30 neighbours to ensure points on ears are not removed for early developmental stages.

Thus, points with less than 30 neighbours were eliminated from further processing.

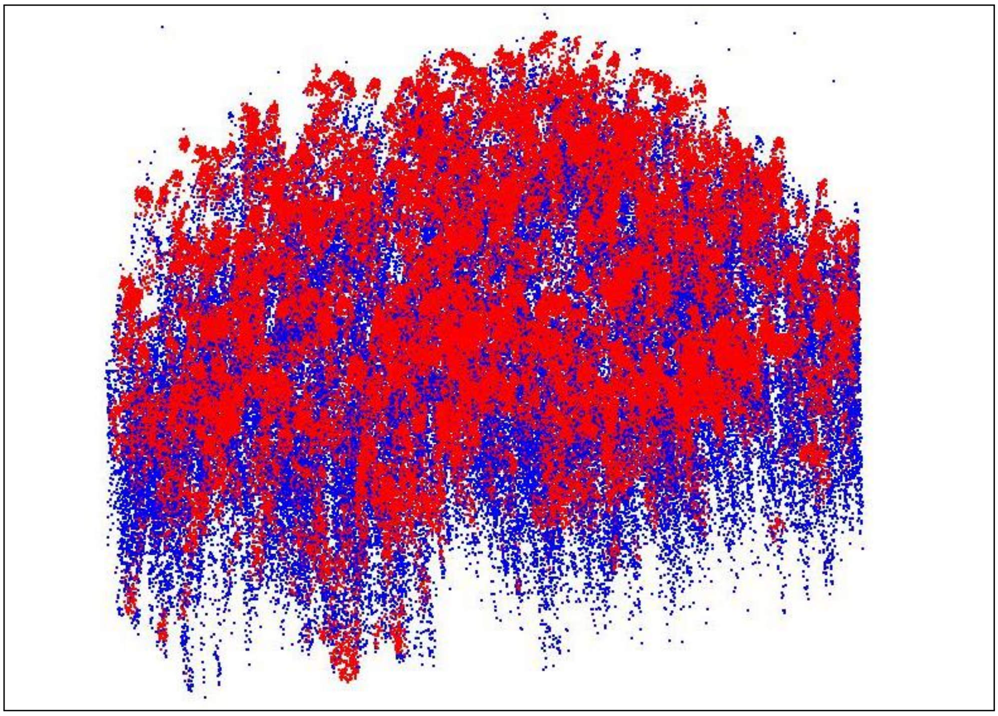


Figure A.1: The points are colour coded based on the number of neighbours present within their local cylindrical neighbourhood. Red = Points with less than 30 neighbours; Blue = Points with more than 30 neighbours (all the ear points fall within this range); (A subset of *Plot C*)