Identification and Location Derivation of Grapevine Features through Point Clouds

BY

DI GAO

A THESIS SUBMITTED IN FULFILMENT OF THE REQUIREMENT FOR THE DEGREE OF MASTER OF ENGINEERING SCIENCE

AT

SCHOOL OF MECHANICAL ENGINEERING

THE UNIVERSITY OF ADELAIDE

APRIL 2014

Contents

Declarations

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I give consent to this copy of my thesis, when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library catalogue and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

Name: Di Gao

Signature:

Date:

Acknowledgement

This is an acknowledgement to those who have given me direction, guidance and who have helped me in my master’s studies

I would like to take this opportunity to express my deepest and most sincere thanks and expressions of gratitude to my supervisors, Dr. Tien-Fu Lu and Dr. Steven Grainger. Their broad knowledge and logical pattern of thought combined with their generous and friendly manner have been of the greatest assistance to me. Their encouragement guidance and mentorship have enabled the presentation of this thesis.

Publications

Conference:

Di Gao, Tien-Fu Lu and Steven Grainger, 'Post Identification and Location Derivation in Vineyards through Point Clouds using Cylinder Extraction and Density Clustering', *2013 IEEE Conference on Robotics, Automation and Mechatronics (RAM)*, 11-14, Nov, 2013, Manila, Philippines.

Journal:

Di Gao, Tien-Fu Lu and Steven Grainger, 'A New Method of Feature Extraction and Location Derivation in Vineyards using Point Clouds', *Applied Engineering in Agriculture*, 30(2): 293-306.

Abstract

An automatic pruning machine is desirable due to the limitations and drawbacks of current labor intensive grapevine pruning methods. Automation mitigates the issue of skilled worker shortages and reduces overall labor cost. To achieve autonomous grapevine pruning accurately and effectively, it is crucial to identify and locate some key features including post, trunk, cordon and cane in order to open/close the cutter and adjust the height of the cutter appropriately. In this thesis, a new method is proposed to automatically identify these features and derive their locations using point clouds. This method combines the advantages of cylinder extraction, density clustering and skeleton extraction for identification purposes. More importantly, it fills the gap of non-uniformed feature extraction in vineyards using point clouds. The results of applying this method to different data sets obtained from vineyards are presented and its effectiveness is demonstrated.

List of Figures

Figures of Chapter One

Figures of Chapter Two

Figures of Chapter Three

Figures of Chapter Four

Figures of Chapter Five

List of Tables

Chapter 1 Introduction

1.1 The Grapevine

The grapevine is a genus of long-lived deciduous vining plants in the flowering plant family vitaceae. It is economically important as the source of fruit and wine production. There are mainly two parts of the grapevine, nonvisible roots and visible tendrils (Figure 1.1). The roots anchor the grapevine to the soil and serve as the conduit where by nutrients and water from the soil are absorbed. Tendrils are used to support the grapevine by clinging to surrounding structures such as a grapevine training system. Tendrils consist of a trunk, shoots and buds. Shoots which sprout from the trunk contain nodes where new leaves, flowers and tendrils can form. At the joint where leaves connect to the shoot are buds that contain the structures that will grow into shoots, tendrils, flowers and leaves of the following years.

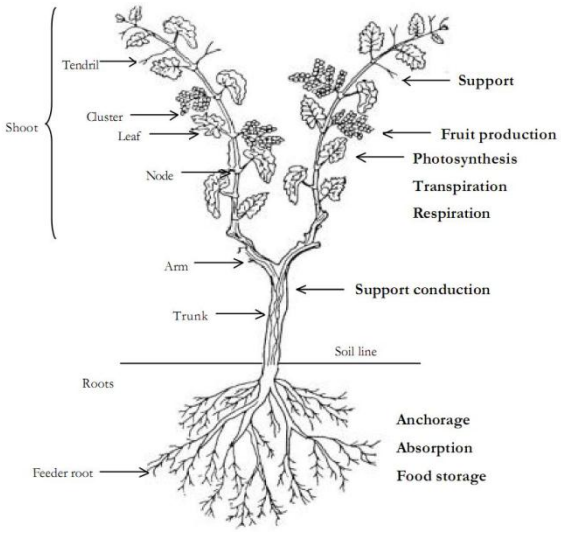


Figure 1.1 Structure of Grapevine
(Tyrin & Barkai 2009)

1.1.1 Annual Growth Cycle of The Grapevine

There are four seasons of a year. The growth season is in spring, when buds swell, leaves grow and shoots arise. During the production season in summer, shoots

produce flowers that develop into grapes. In autumn, grapes are ready for harvest. Then, the leaves turn red or yellow and fall from the grapevine in late autumn. Finally, the bare grapevine is pruned in the dormant season of winter.

1.1.2 Grapevine Training System

Since the grapevine is a kind of vining plant, it requires support to keep it off the ground. Such support is called a grapevine training system or trellis system. The grapevine training system is also aimed to assist canopy management with two aspects. It optimizes sunlight interception to facilitate photosynthesis that affects grape ripening. It also provides the grapevine with better air movement to prevent disease. Another benefit of utilizing different training systems is to facilitate vineyard tasks including applying pesticide, fertilizing sprays, harvesting the grapes, irrigation and the focus of this research, grapevine pruning.

The structure of a training system is shown in Figure 1.2 which is the scene of vineyards in winter before pruning. The post is the stake that wires are attached to. The trunk is the main branch of grapevines. Cordons, which are the two main branches expending from the top of the grapevine trunk, are trained horizontally along wires. Canes are the scattered branches.

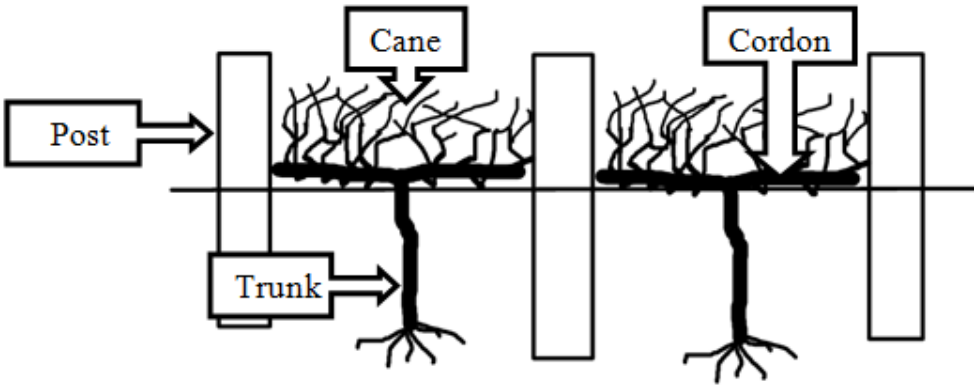


Figure 1.2 Structure of Grapevine Training System

1.2 Grapevine Pruning

1.2.1 Reasons for Grapevine Pruning

Grapevine pruning is the most significant operation in maintaining and determining the vine structure (Bartsch 2010) and it directly influences grapes yield and quality in the following season and beyond (Hoare 2009). Functional and main reasons to prune include:

- It improves the overall health of the grapevine by removing older canes which encourages the grapevine to put energy into new growth and stay young.
- It controls growth and maintains the form of the grapevine which benefits other vineyard operations like harvesting.
- It increases the number and quality of grapes and prevents the spread of diseases in the next year.

1.2.2 Methods of Grapevine Pruning

There are two methods that have been used in grapevine pruning including manual pruning and semi-automatic pruning.

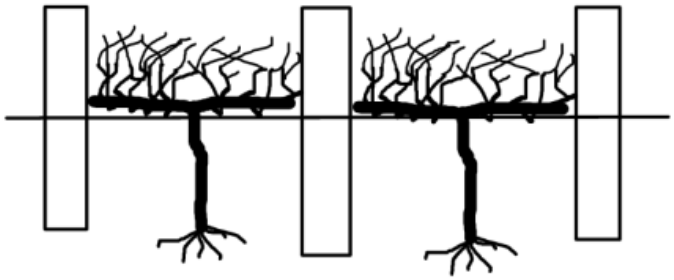
1.2.2.1 Manual Pruning

Manual pruning or hand pruning began with the Romans (Hoare 2009) and has been using trimming tools such as saws, shears and secateurs as shown in Figure 1.3-a. These tools are still widely used in many situations especially in developing countries with a lower labour price. Electrical trimming tools invented such as the rotary saw and electric secateurs (Figure 1.3-b) improve the efficiency of hand pruning.

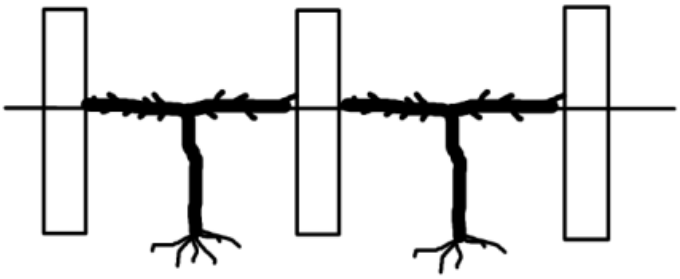
(a) Mechanical Secateurs (b) Electric Secateurs

Figure 1.3 Mechanical and Electrical Secateurs

The pruners perform the pruning operation using these trimming tools to cut down the scattered canes and leave one bud or two bud spurs based on different needs. The grapevine in vineyards before and after pruning through manual pruning is shown in Figure 1.4.



(a) Before Pruning



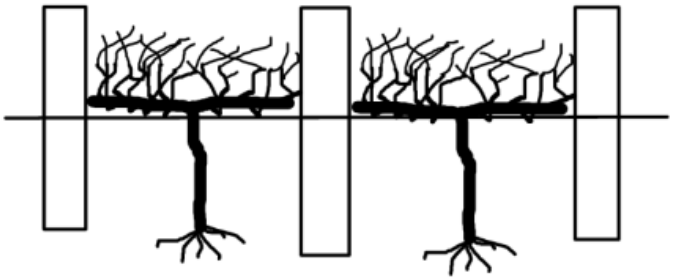
(b) Pruning Complete

Figure 1.4 Before and After Manual Pruning

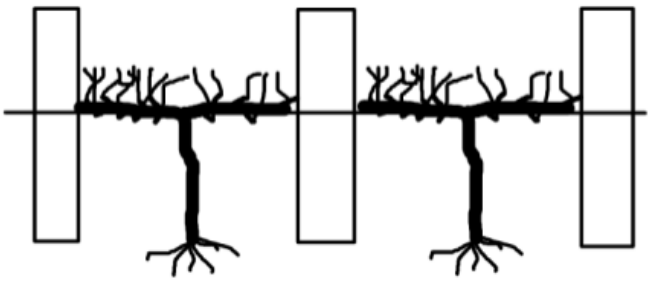
The traditional mechanical tools and the new electrical tools cannot fulfil the needs of grapevine pruning due to the fact that vineyards nowadays are normally very large and there are a large number of grapevines in them. The high cost of labour in developed countries, the skills shortage and need for experienced pruners, and the low efficiency and time consumption of hand pruning are issues still to be properly addressed.

1.2.2.2 Semi-Automatic Pruning

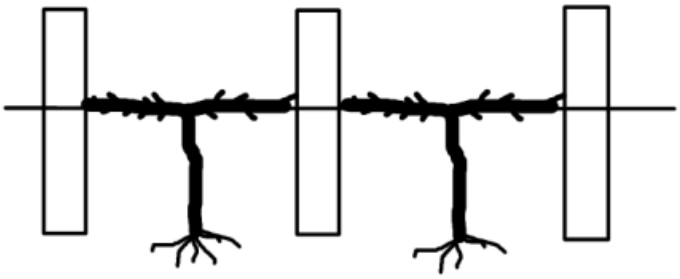
To improve the efficiency of manual pruning and reduce labour cost, a semi-automatic pruning method is applied by dividing pruning into two stages. Stage one is to cut the canes to a certain level by a manually operated pre-pruning machine. Then, stage two goes back to hand pruning. The two stages of the semi-automatic pruning method are illustrated in Figure 1.5.



(a) Before pruning



(b) Stage One



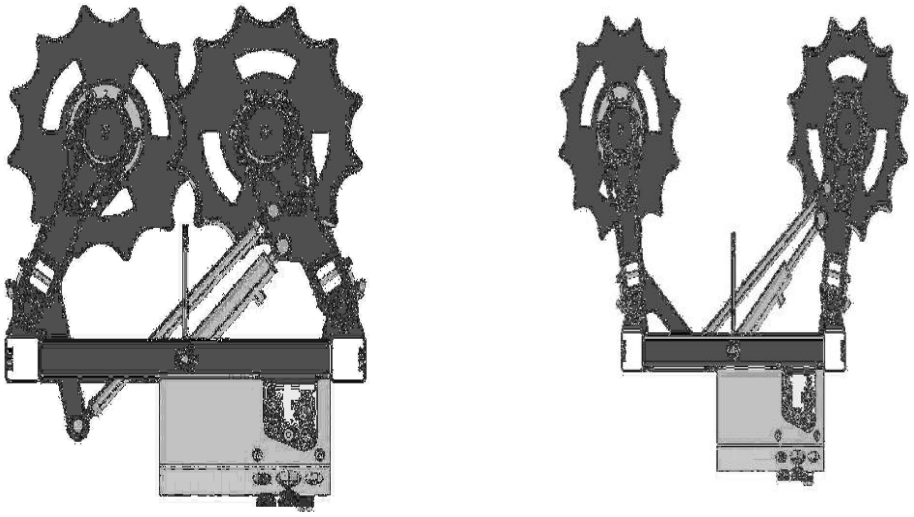
(c) Stage two

Figure 1.5 Two Stages of Semi-Automatic Pruning Method

There are several types of pre-pruning machines which have the similar structure including one tractor and one or two mechanical arms with two cutters. The driver of the machine can adjust and control the position of the mechanical arm and open or close the two cutters to avoid cutting obstacles like posts. Shown in Figure 1.6, it is a commercial product of Pellenc Company for the pre-pruning operation in vineyard called DISCO.

(a) Pre-pruning machine

(b) Two cutters



(c) Cutters closed (d) Cutters open

Figure 1.6 Pre-pruning Machine with One Arm and Two Cutters

Such machines can perform the pre-pruning operation quite fast which indeed improves the overall pruning speed and partly liberates pruners and workers from the hard physical work. However, the accuracy of the operation depends on the experience and capability of the driver who commands and controls the machine. The driver needs to firstly move the tractor in front of one row of grapevines. Then, the mechanical arm is adjusted to a suitable height for the two cutters to prune. After that, the tractor is driven along the row and two cutters cut down the canes. The driver opens the cutters when it is near a post and restores it to closed immediately after passing a post in addition to moving the cutter higher or lower to follow the heights of cordons. During such high concentration operations, the driver feels fatigue which affects the pruning performance and accuracy.

1.3 Motivation

From the discussions above, current grapevine pruning methods still have issues and drawbacks. In terms of manual pruning, the high cost of labour in developed countries, the low efficiency and time consumption of physical work, and the need of experienced pruners are still unsolved even with the invention of electric trimming

tools. On the other hand, the semi-automatic pruning method indeed improves the overall pruning speed and is more automatic than hand pruning. But the pre-pruning machines are driven and manipulated by workers. The accuracy of the pre-pruning highly depends on the experience and capability of the driver who commands and controls the machine. As the driver can easily feel fatigue and tired during the high concentrating driving, the rough pruning of these machines will have negative influences on the grapevine.

Therefore, an automatic pruning machine is desirable in grapevine pruning. This machine will perform the pruning operation autonomously with a minimal or no human intervention. Hence, the accuracy and efficiency of pruning can be improved, the cost of labour can be reduced, and the issue of skills shortage and need of experienced pruners can be resolved.

1.4 Research Purpose

To achieve fully autonomous grapevine pruning, the machine needs to perceive the environment in the vineyard. It needs to identify, recognize and locate key features in the vineyard related to pruning operation which are the post, cordon and cane. There are several challenges to attain the goal. Firstly, the vineyard environment is complex due to the fact that illumination changes at different times of the day and different days as well as in different weather conditions. There are occlusions such as random canes and different backgrounds like other rows of grapevines. Moreover, feature extraction of objects is another challenge. In vineyards, each post is different. The cordons though looking similar are all different as well. As a result, feature recognition required for this research is more difficult than when identifying targets with fixed shapes.

In brief, the research purpose is to investigate and develop methodologies and algorithms to identify and derive the locations of post, trunk, cordon and cane automatically in vineyards.

1.5 Thesis Synopsis

The content of this dissertation is divided into four sections. The first section reviews the applications of machine vision and feature extraction methods of point clouds for vineyard feature extraction purpose (Chapter 2). The second section introduces the two main parts of the new proposed method including segmentation and object identification (Chapter 3 and Chapter 4). The third section provides experiment results by applying the proposed method in order to demonstrate its effectiveness (Chapter 5). The last section of this thesis summarizes the contributions and suggests avenues for future research (Chapter 6). A brief description of each chapter is provided below:

Chapter 2: Literature Review. This chapter firstly reviews applications of machine vision techniques, paying special attention to their applications in vineyard autonomous pruning. It also reviews feature extraction methods of point clouds including basic features, cylinder features and skeleton features.

Chapter 3: Segmentation. This chapter mainly describes the first part of the proposed method with three steps, pre-processing, cylinder feature extraction and density clustering. The crucial concept of the proposed method is introduced as well as the fundamental theories of cylinder feature extraction and density clustering.

Chapter 4: Object Identification and Location Derivation. This chapter mainly describes the second part of the proposed method with four parts, post, trunk, cordon and cane identification. The key concept of skeleton extraction and how to obtain main component from the skeleton extracted are introduced.

Chapter 5: Experimental Work. This chapter focuses on how to analyse and test the effectiveness and performance of the new proposed method for different vineyard scenarios. The experimental setup introduces the range sensors applied, software designed and field configurations. The experiment results show the effectiveness and performance of the new proposed method applied to several data sets in different

situations of vineyards.

Chapter 6: Conclusion. This chapter summarizes what has been discussed in the thesis.

Chapter 2 Literature Review

2.1 Applications of Machine Vision

In order to extract features and derive locations of post, trunk cordon and cane in vineyards, one approach is to use machine vision. It is a digital computer technique which extracts, characterizes and interprets information from visual images of a three dimensional world. As the existing literature on machine vision techniques is vast, works mostly related to this application are reviewed.

In terms of machine vision application in vineyards, it includes weeding robots, autonomous pesticide spray systems, and autonomous pruning machines.

Zhang et al. (2007) developed a new dynamic image measure technique for weeding robots. The technique discriminates the weed and the trunk and measures the root of the trunk. A stereo vision system (Figure 2.1) was applied to capture images of vineyards as input data. The concept of their method is the trunk is in a static status and the weed is in a dynamic status when they are in the wind. The trunk is generally hard while the weed is soft. Hard objects like trunks will almost be static while soft object like weed will be waving all the time. The weed is identified by comparing several images of one scene to find the dynamic field (Figure 2.2). There are two limitations when considering a vineyard scene. Firstly, the post, trunk and cordon are in a static status according to their theory and it is hard to distinguish them using this method. A cane is soft but not as soft as a weed, whether the method can separate canes from other objects still needs further investigation and test. Besides, the illumination condition is not considered and the performance of the method under different lighting is unknown. Igawa et al. (2009) proposed a method of trunk recognition using visual and tact sensing technique. They applied the method of Zhang et al. (2007) to locate the grapevine root. Then, the tact sensing techniques is applied, which is using the multi-link manipulator and can recognize the trunk of

the grapevine by touching it. This method also has the issue that the reliability is uncertain under different illumination condition.

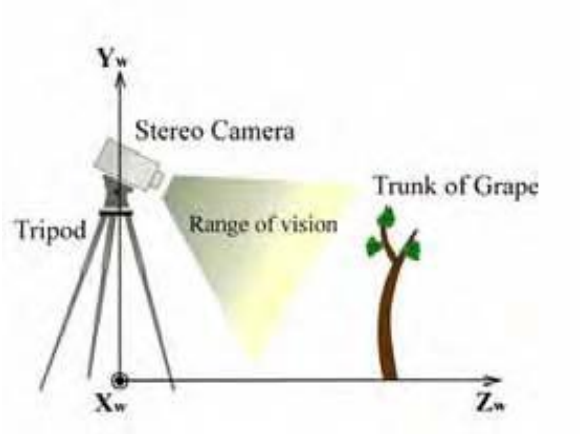
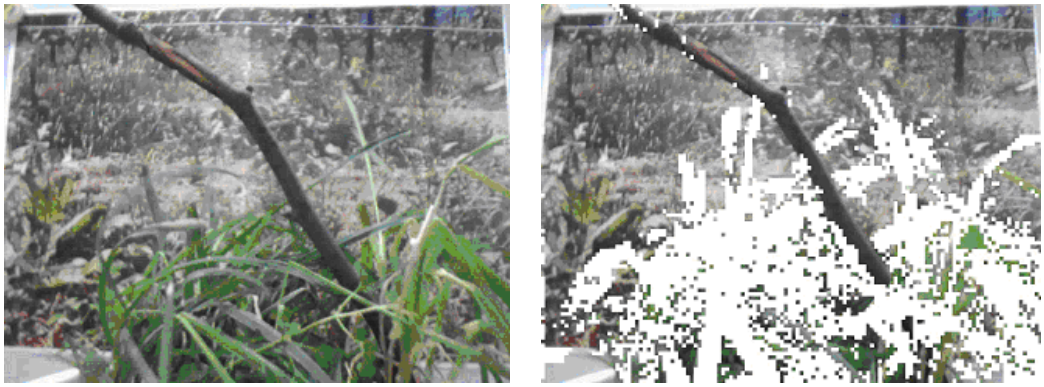


Figure 2.1 Stereo Vision System
(Zhang et al. 2007)



a. Original Image b. Weed Detection

Figure 2.2 Weed and Trunk Separation Method
(Zhang et al. 2007)

In the field of pesticide sprays in vineyards, Braun et al.(2010) proposed a novel visual evaluation technique which can record foliage distribution based on wide range cameras to accomplish the spraying process with high precision. The background was simplified by putting a distinctively coloured canvas behind the vine and changing illumination was mitigated by building a black sunshield around the camera. Whether their method can be applied to identify other objects in vineyards still need further

investigation. Besides, vineyard related operations also need to be performed during night time. The performance of their method operated with fixed lighting conditions at night time is unknown.

Two research efforts are identified in the literature which applied machine vision techniques to extract features in vineyards for autonomous pruning. McFarlane et al.(1997), who focused on the stage of winter pruning, developed image analysis algorithms for locating wires and trunks in long wood grapevines (Figure 2.3). They applied a monocular vision system with a CCD camera to capture original images of vineyards. A powerful flash was used for the camera to eliminate daylight effects. After noise removal including thresholding, size-filtering and segmentation of images, the wires and the trunk are located with their algorithm. Gao & Lu(2006) calculated cutting positions for an autonomous pruning machine using a digital camera (Olympus IR-300) which was also a monocular vision system. They simplified the background by putting a white curtain behind the cordon and canes and using fixed illumination. The colour space of input images is converted from RGB to Black and White. Then, the cutting positions (Figure 2.4) were located based on their new algorithm. Both of the studies found ways of dealing with illumination changes, but the real situation in vineyards is complex and different. From daytime to night time, how to maintain fixed lighting conditions under different weather still needs further investigation.

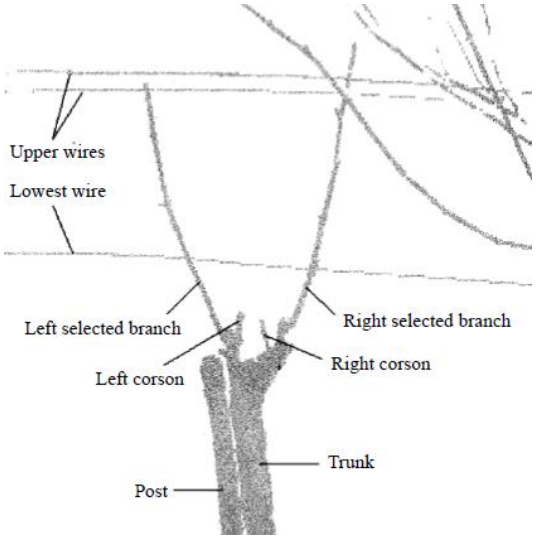


Figure 2.3 Wires and Trunk of Long Wood Grapevine
(McFarlane et. al 1997)

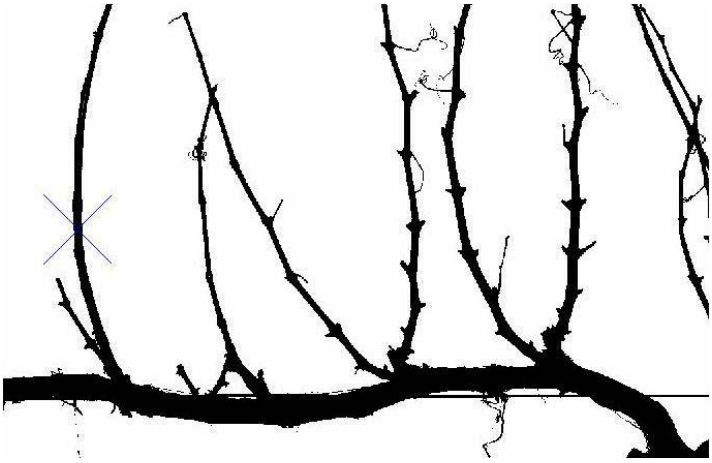


Figure 2.4 Cutting Positions of Canes
(Gao & Lu 2007)

Although the contribution made by these research efforts is significant, there are some common issues. One disadvantage of machine vision is its need for defined and consistent lighting condition (Gümüş, Balaban & Ünlüsayın 2011). These studies are constrained by this limitation and also the need for simplified background which make them impractical to be implemented in real-life vineyard scenarios complicated by rows of vines and operated under very different lighting conditions from time to time. So far, to overcome the changes of data caused by illumination changes perceived by machine vision remains a popular research problem to be properly

addressed. Besides, monocular vision systems cannot provide depth information without knowing the height of the cameras and the distance to the objects. The pruning operation is performed by tractors driven at a quick speed on uneven ground, the accuracy of input data is a big issue.

2.2 Feature Extraction from Point Clouds

Another approach which can avoid such limitations of current machine vision techniques is using point clouds. It is a set of vertices in a three-dimensional coordinate system, which is normally generated by range sensors including laser and stereo vision systems. Point clouds generated by some of the range sensors are not affected by illumination changes, and may be suitable for pruning operations in vineyards around the clock, when needed. Moreover, complex backgrounds like other rows of grapevines, green grasses on the ground and the blue sky can be easily removed using distance threshold filtering. As no literature has been identified in terms of point cloud feature extraction in vineyards, related and relevant literatures are reviewed and detailed in the following.

2.2.1 Basic Features

Basic features of point clouds like lines or sharp edges are widely applied. Weber et. al (2010) proposed a new technique for automatic sharp edge extraction based on Gaussian map clustering (Figure 2.5). Park et. al (2012) presented a new method to extract lines from unstructured point cloud data through multi-scale tensor voting (Figure 2.6). The scene in vineyards however is more complex, because of a combination of natural and artificial products such as trunks, cordons, canes and man-made posts. They often overlap and are always very close to each other. Therefore, there are no specific patterns that can be used to identify these objects using features of lines or sharp edges only.

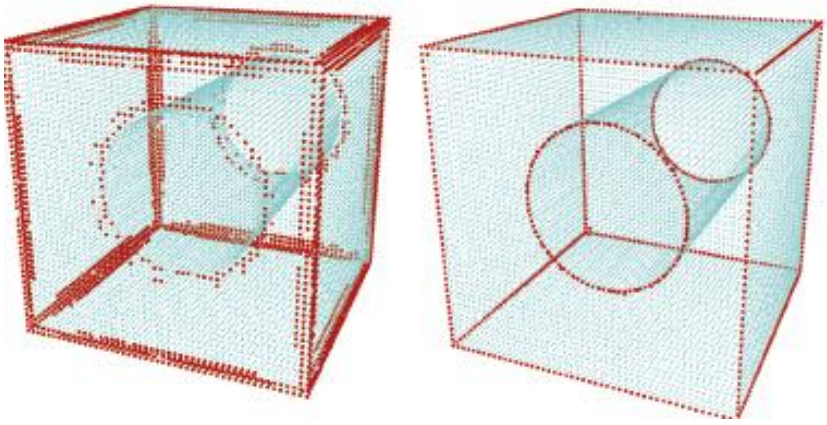


Figure 2.5 Result of Gauss Map Clustering Method Applied to Cub-with-hole Sample

(Weber et. al 2010)

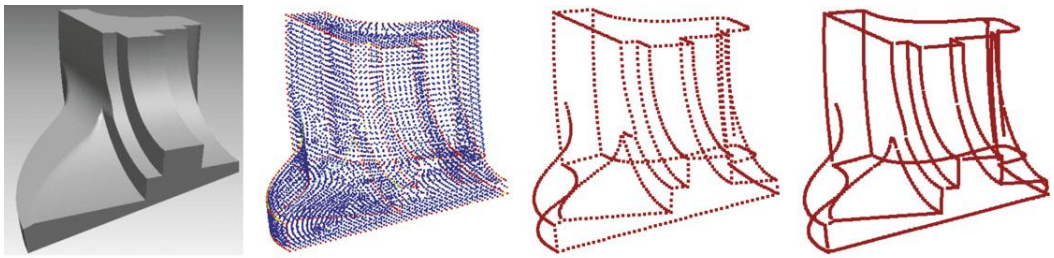


Figure 2.6 Sharp Edge Detection Result for Fandisk Model

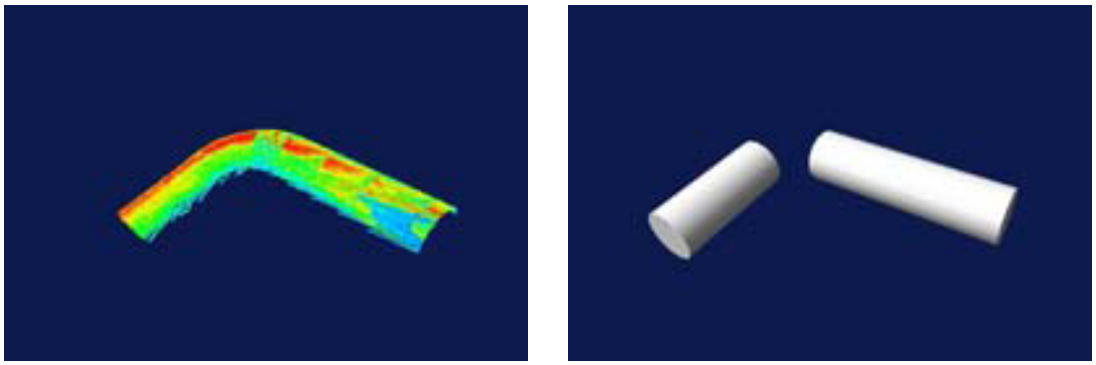
(Park et. al 2012)

2.2.2 Cylinder Features

Since the posts, cordons and canes can be basically described as cylinders, either one cylinder or a series of connected short cylinders, cylinder extraction algorithms and methods are reviewed. There are two common ways to identify and locate cylinders from point clouds. One is to use all the cylinder parameters synchronously (Beder & Förstner 2006; Lukács, Martin & Marshall 1998; Schnabel et al. 2006). However, methods like these are currently impractical because the dimension of the parameter space are too high. The parameter space is at least five dimensions to describe a cylinder, therefore processing time spent on estimating each cylinder is too long to be

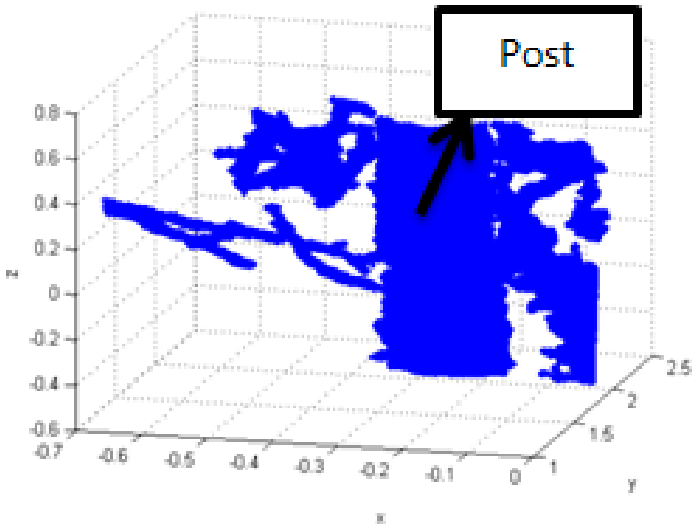
practical. To solve this problem, methods of dividing cylinder extraction into two steps are proposed. Holies & Fischler (1981) proposed a RANSAC-Based approach (RANdom SAmple Consensus) to extract cylinders in range data. The two key steps in this process are fitting ellipses to partial data and fitting lines to sets of three-dimensional points. The limitation is it can only identify and locate cylinders with a known diameter. Lozano-Perez, Grimson & White (1987) developed a method to extract the axis of cylinders based on ellipse fitting. The issue is other parameters of the cylinder such as radius are not extracted and locations are unknown. To not only extract axis but also other parameters of cylinders, Chaperon & Goulette (2001) presented a cylinder extraction method consisting of two steps. Step one estimates orientations of all the cylinders by finding constrained planes in the Gaussian image. The Gaussian image of a cylinder is a great circle, and then the direction of a cylinder can be identified by locating the plane which contains this great circle. Step two is for the cylinder position and radius estimation. Their method solved the issue of high parameter space, although noisy data may cause poor results. Besides, this method may fail to extract cylinders when there are several cylinders of different radii along one orientation (Rabbani & Heuvel 2005). Rabbani & Heuvel (2005) improved the algorithm and solved the problem mentioned using the Hough transform. After using the 2D Hough transform in the first step to find points vote for the possible orientations of cylinders, the radii and positions of the cylinder are calculated and located by circle fitting using the 3D Hough transform with a user-specified radius range (Figure 2.7). However, step two of their method has limitations when considered for the vineyard scenario presented in this research. Firstly, a user-specified radius range cannot be obtained in some situations. Moreover, the results of circle fitting are highly relying on the accuracy of input data. As shown in Figure 2.8-c, there is one post in the middle and several scattered canes around it. Based on the assumption that the post area has been identified, a circle is fit to the post area in order to locate post and calculate its radius. The radius (r1) of the assumed circle fitted is much bigger than the radius (r2) of the actual post. Su & Bethel (2010) improved the first step, orientation estimation, of Rabbani & Heuvel's

(2005) method. They claimed that the run time is decreased and robustness is improved, although the method was applied to two separate cylinders only rather than a scene with several cylinders overlapped and connected. Besides, the radius issue of Rabbani & Heuvel's (2005) method is still unsolved.

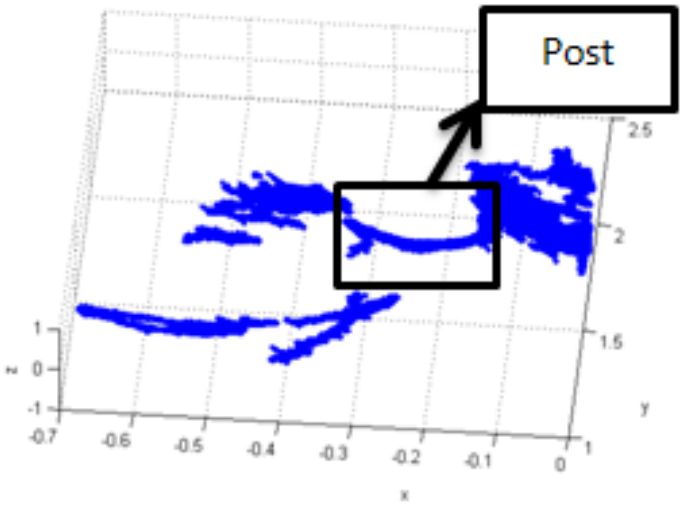


(a) Input data (b) Result of algorithm

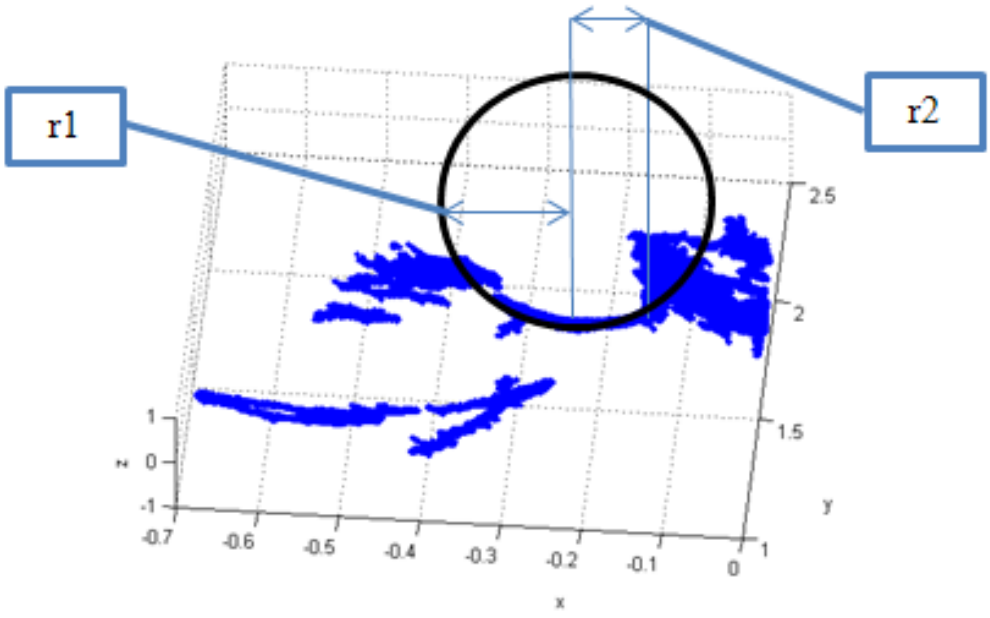
Figure 2.7 Result of Cylinder Extraction based on Hough Transform (Rabbani & Heuvel 2005)



a. Front View



b. Top View



c. Assumption of Circle Fitting

Figure 2.8 Point Clouds of Post in Vineyard

In brief, even if all the cylinders from input points can be extracted and located, the relations among these cylinders remain unknown using the current approaches. For instance, the cordons of grapevines can be described as several cylinders connected with each other. However, there is no way to define a cordon if the relations of these cylinders are unknown. For this reason, methods and approaches that can provide

relations among points and features need to be reviewed, which skeleton extraction is identified to be promising and therefore is a focus.

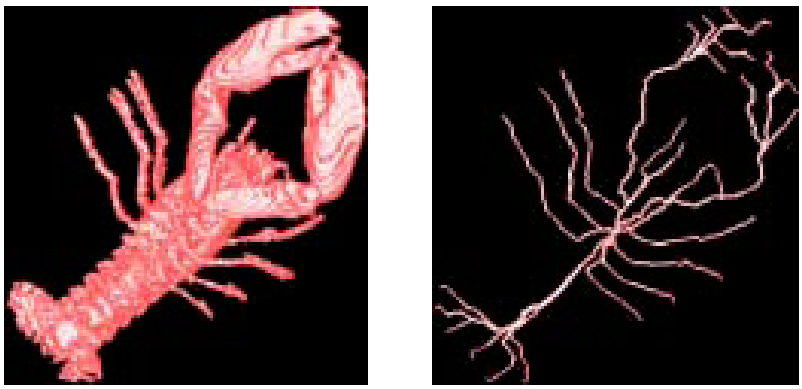
2.2.3 Skeleton Extraction of Point Clouds

2.2.3.1 Simple Structure Skeleton

Simple structure skeleton extraction like artificial products or structures with specific patterns like the human body is not suitable for this research. Because the scene in the vineyard is more complex and no specific pattern can be followed. The skeleton of a grapevine is very similar to that of trees, which is a trunk and several branches. The post can be described as a tree with no branches. Therefore, tree-structure skeleton extraction methods are reviewed.

2.2.3.2 Tree-like Structure Skeleton

Sato et. al (2000) proposed a method called TEASAR (tree-structure extraction algorithm delivering skeletons that are accurate and robust) to find skeletons accurately and rapidly. A multi-stage Skeletonization method for tree-like volumes has also been developed (Chen et. al 2000). Both of the methods were tested for tree-like targets including lobster (Figure 2.9), ribs and colon.



(a) Lobster (b) Lobster Skeleton

Figure 2.9 Tree-like skeleton extraction - Lobster

2.2.3.3 Tree Structure Skeleton

There are also methods directly applied to trees. Gorte & Pfeifer (2004) developed a method to extract a skeleton of laser-scanned points using 3D mathematical morphology. The key concept of this method is a 3-dimensional raster domain which is a discrete space with elements called voxels. They applied voxels to analyze and extract skeletons of trees instead of points. The drawback of Gorte & Pfeifer's (2004) method is that a large number of algorithm parameters need to be controlled (Bucksch & Lindenbergh 2008). For example, the spatial resolution of the 3D raster domain which denotes the size of voxels is selected based on the density of laser points and processing time. It may become impracticable when the resolution is chosen too fine (Gorte & Pfeifer 2004). Besides, there is a tendency of producing cycles in the extracted skeleton. Gorte (2006) developed a Dijkstra skeletonization method based in 3D raster domain. It automatically identifies the numbers, lengths and thickness of tree branches. The main advantage of this method is it follows the original data closely and does not need pre-processing. The methods of Gorte & Pfeifer (2004) and Gorte (2006) share one common drawback that both of them are restricted to terrestrial laser scanners. The input points of trees are scanned from all sides by locating the scanner at several viewpoints (normally 4) around the target. The effectiveness of their methods on points captured by other range sensors with only one side scan needs further investigation.

Bucksch & Lindenbergh (2008) presented a Skeletonization algorithm called CAMPINO (collapsing and merging procedures in octree-graphs). CAMPINO is able to extract skeletons through point clouds captured from one or multiple viewpoints as shown in Figure 2.10. Bucksch, Lindenbergh & Menenti(2009) reported a new algorithm aiming at the skeletonization of a laser scanning point clouds. Both of these two methods are based on the Octree concept, which generate an Octree and extract a graph from points in the Octree cells. The drawback is the use of a preset partitioning resolution, which makes the result unreliable when the branches

have different densities in different parts of the tree (Livny et al. 2010).

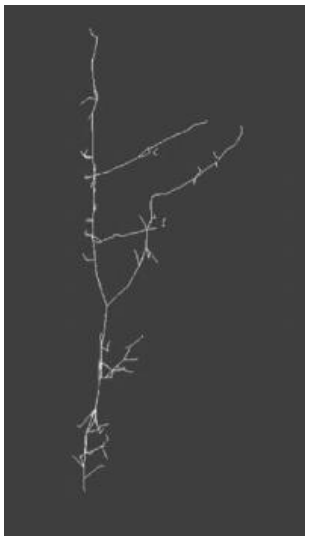
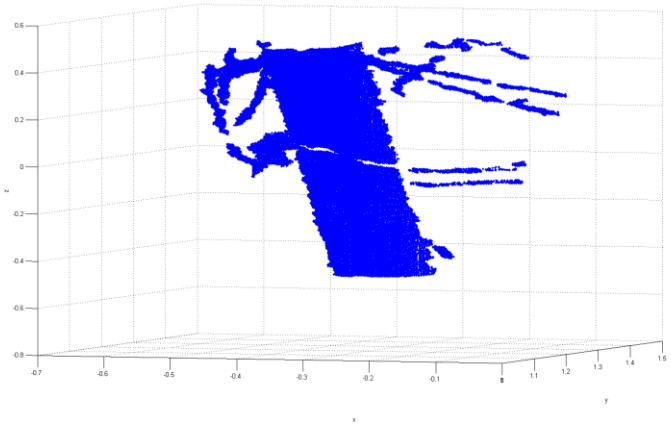
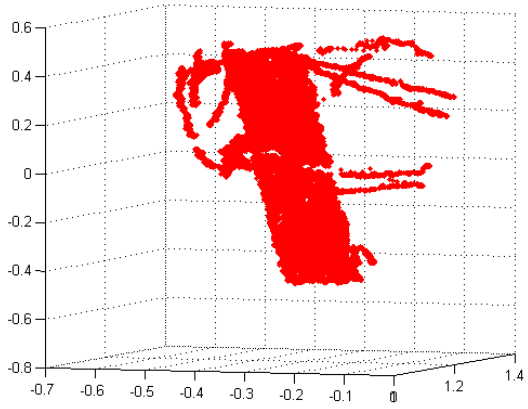


Figure 2.10 Collapsing and Merging Procedures in Octree-graphs (CAMPINO) (Bucksch et. al 2008)

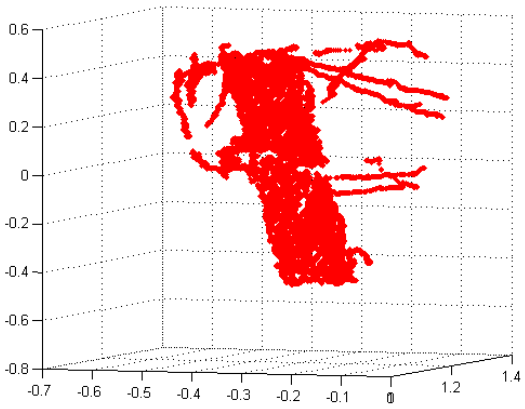
There is a common issue of previous methods that are only suitable for one tree in the scene. To solve these issues, Linvy et al. (2010) proposed an automatic tree reconstruction algorithm using a global optimization method which is more robust to noise, ununiformed point density, missing data and multi-trees situation. The result shows not only the skeleton is extracted but also the geometry aspects of the tree. Although the method of Linvy et al. (2010) is able to extract skeletons in the multi-trees situation, there is a certain distance among trees which makes it easy to pre-processing the points to separate trees. The scene in vineyards is more complex and different. One post and one grapevine can be very close or even overlapped with each other. How to extract the skeletons of post and grapevine in such an environment is still open to be investigated. Besides, the input points of their method are sparse. When it comes to situation like vineyards, the high density area such as post can lead to poor skeleton results based on the understanding of the paper and test results. It is shown in Figure 2.11, the algorithm can generate a good skeleton result when applied to the canes combined with sparse points. However, the skeleton of the post area is still unclear.



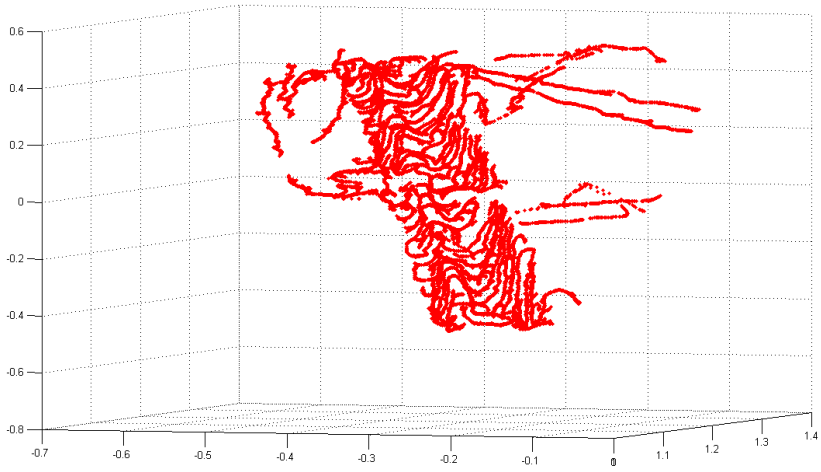
a. Post with several canes



b. Iteration one



c. Iteration Four



d. Iteration Eight

Figure 2.11 Test Result of Skeleton Extraction Method (Linvy et al. 2010) Applied to Data of This Thesis

2.2.3.4Curve Skeleton

As current tree skeleton methods have limitations in either the number of algorithm parameters, the restriction to certain laser scanners or the poor result in vineyards, curve skeleton methods are reviewed. Curve skeleton methods do not need a large number of parameters, do not restrict certain laser scanners and can be applied to different densities of point clouds, which is suitable for the vineyard scenarios. Most of the existing works on curve skeleton extraction operate on complete surface models (Cao et al. 2010). Point cloud data captured in vineyards is different. The input is three dimensional points with missing data. Cao et al. (2010) solved this issue and extracted the skeleton from point cloud data via Laplacian-Based Contraction. Their approach is robust to noise and can handle moderate amounts of missing data which reach the requirement of vineyards. However, there is no literature identified to extract features of post, trunk, cordon and cane even if their skeletons can be extracted.

2.3 Summary

Based on the reviewed literature, the followings are concluded:

Machine vision using cameras is currently not suitable for automatic pruning as performance suffers due to serious illumination changes and the need for simplified backgrounds. Point cloud data generated by some of the range sensors are not affected by lighting changes. Moreover, complex backgrounds like other rows of grapevines, green grasses on the ground and the blue sky can be easily removed using distance threshold filtering. As a result, this research adopted point clouds as input data.

No literature has been identified in the area of extracting natural and artificial features in vineyards through the use of point clouds. Therefore, related areas have been reviewed including cylinder extraction and skeleton extraction. Current cylinder extraction methods have limitations when applied to vineyards. The noisy input points of the vineyard may cause poor results. It also may fail to extract cylinders when there are several cylinders of different radii along one orientation like scattered canes. Besides, parameters like radii of cylinders cannot be obtained, which is essential for some methods. A new or improved method is needed to extract cylinders in vineyards. On the other hand, current tree skeleton methods still have limitations in terms of feature extraction in vineyards. More importantly, even if a way of extracting the skeletons of these four objects can be found or proposed, there is no literature identified to extract the features of posts, trunks, cordons and canes based on only skeletons.

2.4 Gap

No research has been identified in the current literature directly related to post, trunk, cordon and cane identification and location derivation through point clouds. This research is therefore proposed to advance knowledge in this corresponding area. According to the review of related research and analysis of limitations, the gap is

about algorithm(s)/method(s) that can be used to identify post, cordon and canes automatically in the vineyard.

2.5 Aims and Objectives

The aim of this research is to research and develop required algorithm(s)/method(s) to automatically extract features of post, trunk, cordon and cane from noisy 3D point clouds data.

To achieve the aim, the following objectives need to be accomplished:

- Pre-processing input raw point cloud data to remove noise and simplify the scene
- Extracting features of post that is made by artificial wood products
- Extracting the features of trunks, cordons and canes that are in the shape of natural plants
- Identifying post, trunk, cordon and cane
- Deriving locations of post, trunk, cordon and cane

Chapter 3 Segmentation

There are at least four features in vineyards directly related to pruning operations, which are the posts, trunks, cordons and canes. There are also unrelated features, such as the solid ground surface and other rows of grapevines. The post and grapevine in vineyards are normally intersected and occluded with each other. The intersection and occlusion between them normally makes the vineyard scene very complex, therefore segmentation is necessary for identification purpose.

3.1 Introduction

Segmentation is a process of dividing input data into several disjoint areas that maintain the unique and homogeneous features from surrounding. It is understood as a labelling process to classify input data according to the region it belongs to (Josep & José 2008). This process is believed to be able to divide the captured vineyard points into several clusters, which isolates the objects relating to the pruning operation and reduces the complexity of the vineyard scene.

The crucial concept of the segmentation method for this study is to divide the input points into several clusters using two features synchronously, the features of cylinder and density. The post in this study is made by artificial wood products, which can be described as one regular cylinder. Each grapevine consists of one trunk, two cordons and several canes, which can be described as the connected combination of several cylinders. However, applying current cylinder extraction methods has one issue that requires the radii of trunk, cordon and cane which are unknown as they vary depending on the ages of vines and the settings of different vineyards. To tackle this limitation, the feature of point cloud density is applied. It is observed that the point density of post, cordon, trunk and cane are different from each other. Therefore, a density clustering method is likely able to extract the areas of these four features even in the presence of noise. It divides the input point clouds data into several clusters; the

cluster that has a certain number of points which meets the point threshold of a post is labelled as post region ; the clusters has a certain number of points which meets the point threshold of a trunk is labelled as a trunk region; remaining thus contain cordons and canes.

There are three steps to the proposed segmentation method including pre-processing, cylinder feature extraction and density clustering as shown in the flow chart (Figure 3.1).

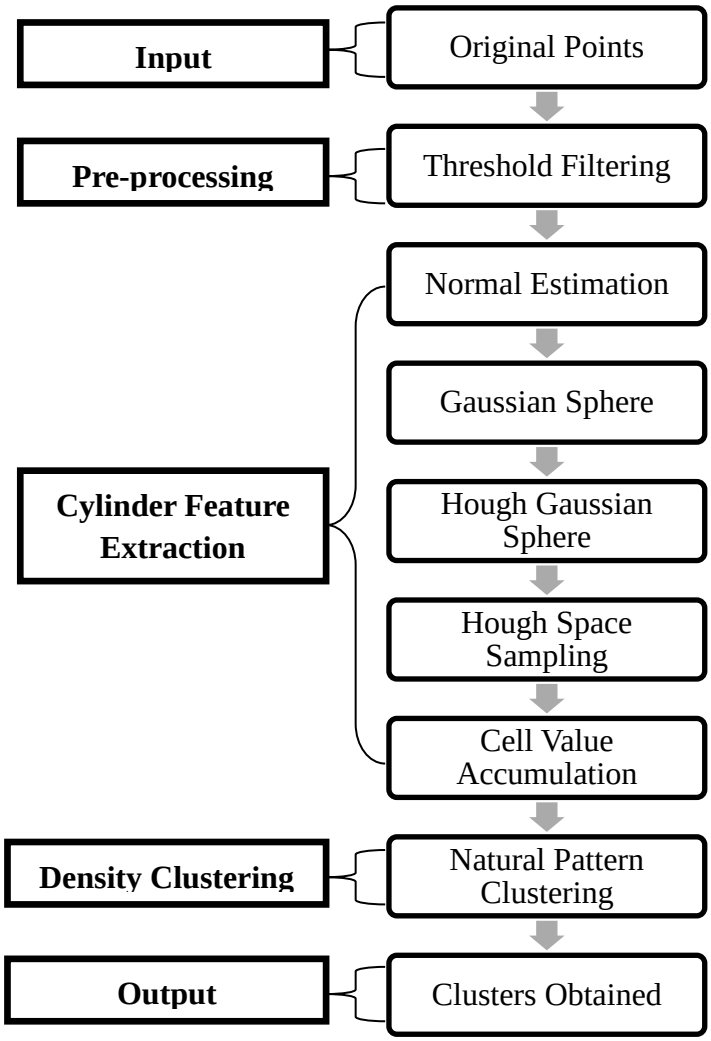


Figure 3.1 Flow Chart of Segmentation

Step one: Pre-processing (section 3.2) is performed in order to remove unwanted and noisy points by threshold filtering.

Step Two: Cylinder feature extraction (section 3.3) is performed in order to make the density features of post and trunk more notable than cordon and cane.

Step Three: Density Clustering (section 3.4) is performed in order to divide all points into clusters.

Output: Clusters obtained from step 3 is the outcome of **Segmentation** (Chapter 3) as well as the input of **Object identification and location derivation** (Chapter 4).

3.2 Pre-processing

The height of the range sensor and its distance to targeted objects can be easily obtained. The distance between two rows of grapevines can also be measured in advance. Therefore, a distance threshold can be directly applied to remove unrelated points, such as ground surface and other rows of grapevines. The ability of the range sensor employed in this study can capture only one row of grapevines, therefore the unrelated points belong to solid ground.

This is one benefit of using point clouds as input rather than images captured through machine vision techniques, the pre-processing is performed easily and precisely.

3.3 Cylinder Feature Extraction

In vineyards, the four features relating to pruning operations are post, trunk, cordon and cane. They can be described as cylinders, either one cylinder like post or several cylinders connected and intersected with each other. Therefore, cylinder feature extraction is performed here.

3.3.1 Cylinder Definition and Parameterization

The definition of a cylinder is described as following:

- The surface generated by a straight line intersecting and moving along a closed plane curve, while remaining parallel to a fixed straight line that is not on or parallel to the plane of the directrix (directrix is a fixed line used in the description of a curve or surface.)

- The portion of such a surface bounded by two parallel planes and the regions of the planes bounded by the surface.

- A solid bounded by two parallel planes and such a surface, especially such a surface having a circle as its directrix.

It needs a minimum of five parameters to describe a cylinder, which means a cylinder has five degrees of freedom in parameter space. There are different parameterizations to represent a cylinder, such as seven parameters with two constraints as used by Lukacs et. al (1998). In this thesis, a five parameters model proposed by Rabbani & Heuvel (2005) is applied as shown in Figure 3.2. This model uses a minimum of five parameters with no constraints that is best suited for cylinder feature extraction. $(\theta, \emptyset)$ gives the axis direction $n = (\cos\theta \sin\emptyset, \sin\theta \cos\emptyset, \cos\emptyset)$ in spherical coordinates. θ is the angle between direction n and axis x and $\emptyset$ is the angle between direction n and axis y. r is the cylinder radius. $P(u, v)$ represents the position in terms of u and v which is along with axial direction.

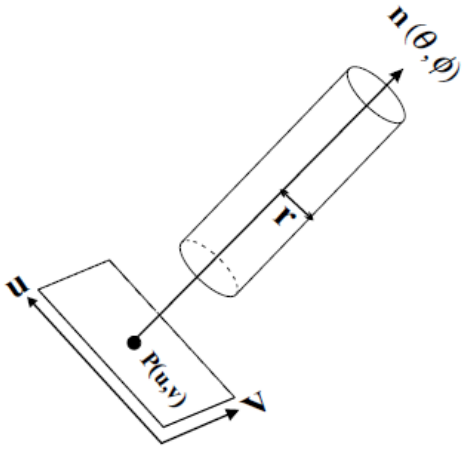


Figure 3.2 The Five Parameters Cylinder Model
(Rabbani & Heuvel 2005)

3.3.2 Hough Transform

Hough transform theory is used in this study in order to extract cylinder features, therefore an introduction of the Hough transform is described as follows. The Hough transform is a feature extraction technique used in image analysis, computer vision, and digital image processing (Shapiro et.al 2001). The purpose is to identify imperfect instances of objects within a certain class of shapes by a voting procedure. This voting procedure is carried out in a parameter space, from which object candidates are obtained as local maxima in an accumulator space that is explicitly constructed by the algorithm for computing the Hough transform.

The Hough transform was first introduced as a method of detecting point's patterns from image data. It examines each point and finds all possible parameters of a specific pattern model. The parameters are collected in a properly defined parameter space, also called Hough space. In the Hough space, the data patterns complied with the specific model can be determined via cluster identification. In practice, a complex pattern detection task is transformed to a more manageable peak detection task in the Hough space. 2D and 3D Hough transform are widely applied to extract features from complex data set.

3.3.2.1 2D Hough Transform

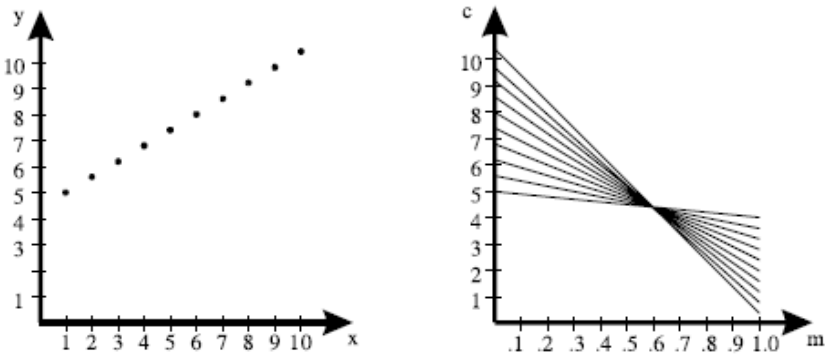
An example of application of 2D Hough transform is introduced here to demonstrate the effectiveness and performance of this technique. The goal is to identify a set of collinear points using a 2D Hough transform in a binary image. In this image, the pixels are wanted when the value is 1. A set of wanted image points (x, y) are collinear when they meet equation:

$$\mathrm{y} - m'x - c' = 0 \qquad (3.1)$$

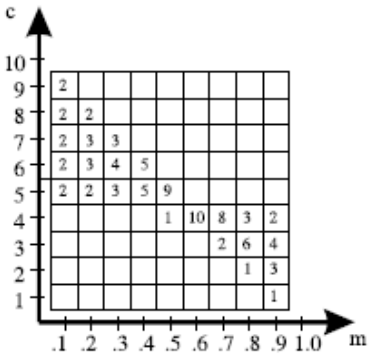
, where m' and c' are the parameters that define slope and intersect of the considered line (Figure 3.3-a). The equation represents a mapping from the image space to the parameter space, a one to many mapping. All the parameters (m, c), that represents the lines passing through a given point (x', y'), can be identified. Such lines satisfy the equation.

$$y' - x'm - c = 0 \qquad (3.2).$$

Therefore, each point (x, y) in the image represents a line in the parameter space defined by (m, c) as shown in Figure 3.3-b. It also means a set of collinear image points has their corresponding lines in the parameter space or Hough space. These lines intersect with each other in one point. The coordinate of this point is the parameter of the line in image plane wanted to be identified. In order to extract this point, an accumulation array or Hough matrix is applied, in where the votes are counted. The sum of votes accumulated in one cell indicates the relative likelihood of lines described by parameters within the corresponding parameter cell (Figure 3.3-c).



(a) Data Space (b) Hough Space



(c) Accumulation Array

Figure 3.3 Hough Transform of Line Detection
(Sarti & Tubaro 2002)

3.3.2.2 3D Hough Transform

The definition of the 3D Hough transform can be described as the natural extension of the 2D Hough transform. In the data space, there is a plane defined in equation (3.3):

$$a'x + b'y + c'z + 1 = 0 \qquad (3.3)$$

In the parameter space, the plane defined represents a point of coordinates (a', b', c'). This means a plane in the data space is represented by a point in the parameter space. On the other hand, if there is a given point p of coordinate (x', y', z') in the data space, a plane in the parameter space passing through it is defined by the coordinate

of this point. The plane in the parameter space is defined in equation (3.4) with parameters of (x', y', z') and variables of (a, b, c):

$$ax' + by' + cz' + 1 = 0 \qquad (3.4)$$

The relationship between data space and parameter space can be demonstrated more obviously by defining a plane in the vector form (equation (3.5)):

$$[a\ b\ c\ 1] \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = 0 \qquad (3.5)$$

There is a severe drawback of the above plane parameterisation (Sarti & Tubaro 2002). The parameter space turns out to be unbounded even if the data space is limited (equation (3.6)). The parameter (a, b, c) can range anywhere from $-\infty$ to $+\infty$.

$$x_{min} \leq x \leq x_{max}, y_{min} \leq y \leq y_{max}, z_{min} \leq z \leq z_{max} \qquad (3.6)$$

Alternative parameterisations are considered in order to solve this issue. One of them is directly derived from the 2D Hough transform. A plane is defined in equation (3.7).

$$x \cos\alpha + y \cos\beta + z \cos\gamma = d \qquad (3.7)$$

α, β and γ are the angles between the plane and the data axis (x, y, z). Besides, the angles are bounded to satisfy the constraint:

$$(\cos\alpha)^2 + (\cos\beta)^2 + (\cos\gamma)^2 = 1 \qquad (3.8)$$

Therefore, the plane parameterisation becomes (α, β, d) which is more suitable for real case application (Sarti & Tubaro 2002).

3.3.3 Orientation Estimation of Cylinders

Hough transform based methods have been applied to solve the problems of outliers

and multiple instances (Hough, 1962). In terms of range data, they are applied to tackle building reconstruction problems by locating planar surfaces in 2.5D data captured from an air-borne laser scanner (Vosselman & Dijkman 2001). Besides, they have been used for characterization of planar fractures from 3D data (Sarti & Tubaro 2002). In this paper, the Hough transform is applied to extract one feature of the cylinder, orientation.

The orientation estimation method (Rabbani & Heuvel 2005) which extracts cylinder orientations using the Hough transform is adopted. The concept of their algorithm is based on the fact that the normal of a cylinder generates a great circle on a Gaussian Sphere (Carmo 1976). This great circle is the result of the intersection of the unit sphere with a plane passing through the origin. The orientation of the cylinder equals to the normal vector of the plane. Once the plane is located, the orientation of the corresponding cylinder can be determined. Points vote for this plane in a Gaussian Sphere are the possible points of the cylinder.

The method of orientation estimation is described as follows:

3.3.3.1 Normal Estimation and Gaussian Sphere

Normal vectors of all points are calculated using KNN (k-nearest neighbours) (Hoppe et al. 1992). Firstly, k nearest neighbours of each point is searched. Then, the normal vectors are estimated by eigen-analysis of their covariance matrix. This normal estimation method is more suited to unstructured point clouds (Rabbani & Heuvel 2005), such as points captured from vineyards. As the normal of all points is obtained, the Gaussian sphere is generated by the normal vectors. As illustrated in Figure 3.4, each point of the Gaussian sphere represents an input point.

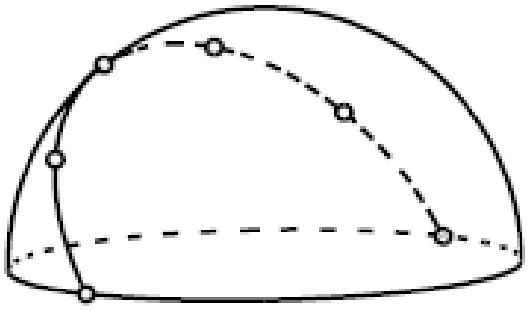


Figure 3.4 Gaussian Sphere

3.3.3.2 Hough Gaussian Sphere

As mentioned, the normal vectors of a cylinder form a great circle on the Gaussian sphere. This great circle is the outcome of the intersection of the unit sphere with a plane passing through the origin. Therefore, the goal is to locate this plane. In order to find planes, the Hough transform requires a three dimensional Hough space (Sarti & Tubaro 2002). There are two parameters including the direction of the plane normal expressed in spherical coordinates and the distance of the plane from the origin. The second parameter which is the distance can be removed because the plane must pass through the origin. Therefore, 2D Hough space is enough for the purpose. Besides, the plane intersects with the unit sphere, which means each point on the Gaussian sphere generates a circle in the Hough space.

In order to generate all the great circles, spherical coordinates is selected which has two variables θ and φ , and one constant r equals to 1 as shown in Figure 3.5.

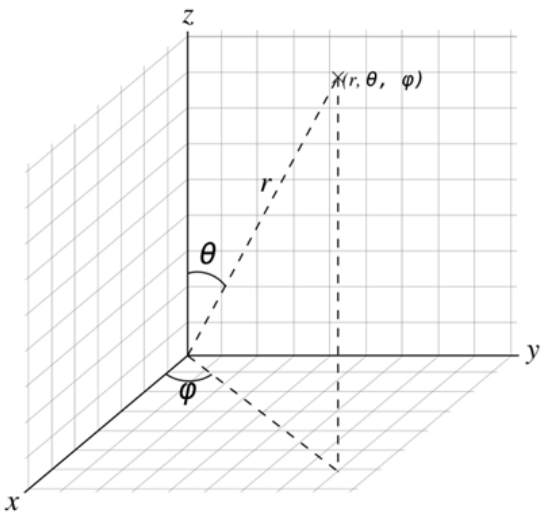


Figure 3.5 Spherical Coordinates

A standard circle is defined in equation (3.9) and has its normal $n_s = (0\ 0\ 1)^T$. The normal of all points is defined in equation (3.10). Then, the great circle of each point is obtained by transferring a standard circle to the direction of all the normal using reflection matrix of Householder. The reflection matrix M is defined in equation (3.11), where I is an identity matrix and b is defined in equation (3.12).

$$\text{xc} = \cos t\,, yc = \sin t\,, z = 0, 0 \leq t \leq 2\pi \qquad (3.9)$$

$$\text{n} = (\cos\varphi \sin\theta \quad \sin\varphi \sin\theta \quad \cos\theta)^T \qquad (3.10)$$

$$\text{M} = \text{I} - 2bb^T \qquad (3.11)$$

$$\text{b} = \frac{n_s - n}{\| n_s - n \|} \qquad (3.12)$$

The radius of the great circles is 1 and their centre is the coordinate origin, which generates another Gaussian sphere called the Hough Gaussian sphere (Figure 3.6).

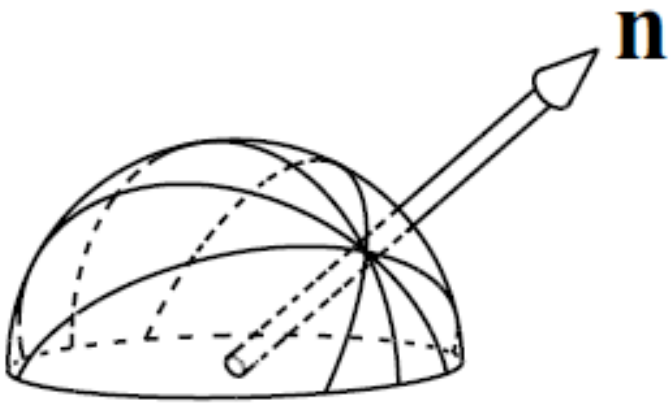


Figure 3.6 Hough Gaussian Sphere

3.3.3.3 Sampled Hough Space

The intersection of these great circles gives a set of values which can be used to estimate the orientation of the cylinders. It needs to calculate the number of the great circles intersected with each other to obtain the values. However, the circle is not continuous round line but discrete points. In order to calculate the values, a sampled Hough space is necessary where the surface of the Hough Gaussian sphere needs to be divided into a set number of cells.

All sampled cells do not have the same size when using a regular quantization in θ and φ (Lutton, Maitre & Lopez-Krahe 1994). Therefore, the approximate uniform sampling method for quantization of Hough space proposed by Lutton, Maitre & Lopez-Krahe (1994) is applied here for better accumulation result. One cell defined by $\Delta\theta$ and $\Delta\varphi$ has a surface area ΔS equal to $\Delta\theta\,[\cos\varphi - \cos(\varphi + \Delta\varphi)]$. N is the number of layers for φ between 0 and $\frac{\pi}{2}$. $\Delta\varphi$ and $\Delta\theta(n)$ referring to nth layer are defined in equation (3.13) and equation (3.14).

$$\Delta\varphi = \pi/2N \tag{3.13}$$

$$\Delta\theta(n) = \frac{\Delta S}{\Delta\theta} \cdot \frac{1}{\cos(n-1) - \cos n} \tag{3.14}$$

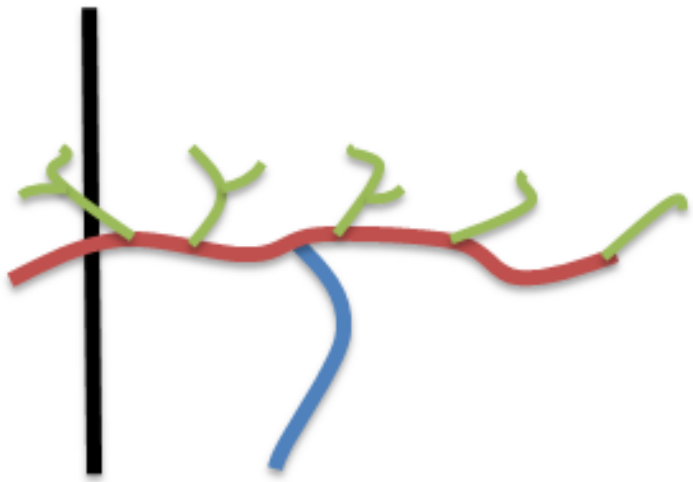
3.3.3.4 Cell Value Accumulation

As the points on Hough Gaussian sphere is known and the sampled Hough space is obtained, the accumulation procedure can be performed. For each cell of sampled Hough space, there is a certain value of number to indicate how many points or great circles pass through it.

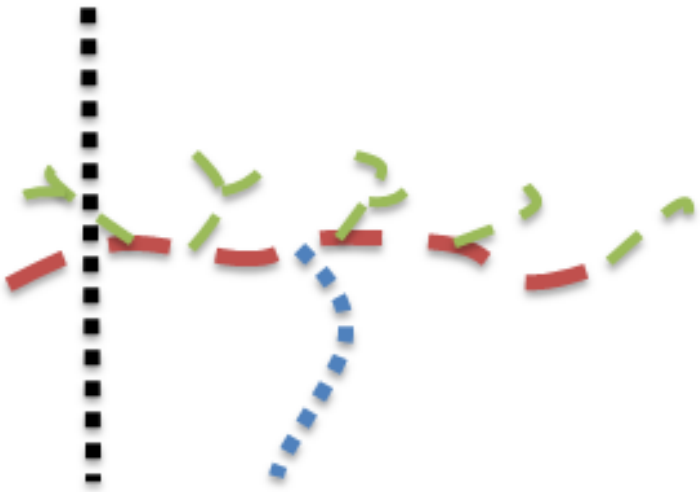
Posts are the biggest and tallest objects in vineyards, which mean the cylinder that represents a post has the highest number of points. Therefore, the cell with the highest accumulator value contains most of the post points.

The points vote for the cell has the highest accumulator value can be easily found. Those are the parts of the points of the post and possible points of the most notable cylinders in the input data.

This step makes the density of post and trunk more notable than other parts in vineyards. As shown in Figure 3.7, more of points belonging to the post and trunk are remained after the extraction than the points of cordons and canes. As the density feature is notably enough, the next step is to segment input data into several clusters using density clustering.



a. Input



b. Cylinder Feature Extraction

Figure 3.7 The Results of Cylinder Feature Extraction

(The colour of post is illustrated in black, trunk is in blue, cordon is in red and cane is in green.)

3.4 Density Clustering

3.4.1 Clustering Methods

Clustering assigns a set of objects into groups so that the objects in the same cluster are more similar to each other than to those in the other clusters.

There are at least five categories of clustering algorithms including Hierarchical clustering, Partitioning clustering, Grid-based clustering, Model-based clustering and Density-based clustering. Hierarchical and Partitioning clustering are not designed for any shapes of cluster but only for globular clusters (Bramer 2007) and therefore not suitable for the project. In the case of Grid-based clustering, boundaries of clusters are either horizontal or vertical (Han & Kamber 2006), but the boundaries of clusters in a vineyard are random because most of the points are from natural grown grapevine. Model-based clustering needs to assign models for each cluster (Han & Kamber 2006).

Objects in vineyards on the other hand are normally natural grown grapevines which is hard to find a specific model for extraction purpose. Besides, the methods aforementioned require an a priori knowledge about the number of clusters contained in a given data set. For natural clusters like vineyards, this number is very hard to obtain and to be certain. Therefore, these methods are not suitable for the needs of the project. Only Density-based clustering methods are suitable for the complex scene in vineyard because it is designed for any shapes of clusters with noise (Ester et al. 1996). Density clustering utilizes a local cluster criterion. Clusters are defined as regions in the data space, where the objects are dense and remain separated from one another by low density areas. It can deal with arbitrary shaped clusters and significantly more efficient for large data sets (Daszykowski, Walczak & Massart 2001).

3.4.2 Density Clustering

NP (natural pattern) clustering method (Daszykowski, Walczak & Massart 2001) is applied here, which is a density-based unsupervised clustering approach developed based on DBSCAN (density-based spatial clustering of applications with noise) (Ester et al. 1996). It adopts the advantage of DBSCAN that is designed for any shapes of clusters with noise. Besides, it works with only one input parameter that is the minimal number of objects considered to be a cluster, which avoids the poor result of DBSCAN when the neighbourhood radius is selected improperly. There are two steps of NP, one is based on the DBSCAN and the other is to optimize the parameter of neighbourhood radius.

3.4.2.1 DBSCAN

DBSCAN is a single scan clustering method relying on a density based clusters notation. The single scan methods structure is described as following:

for (each object i)

if (i is not a member of a given cluster)

create a new cluster C

while (the neighbouring objects satisfy the cluster condition)

add them to C

end

end

end

The main concept of density based clustering is built on the assumption that within a given radius ε, the neighbourhood of each object belonging to a cluster contains at least a minimum number (minpts) of objects. Besides, the density of the neighbourhood exceeds a certain threshold.

The core objects, border objects and outliers are defined as following:

- The ith object is a core object, if in its neighbourhood of radius ε there are minpts objects;
- The ith object is a border object, if it belongs to the neighbourhood of any core object, itself not being a core object;
- The outlier is defined as an object, with the neighbourhood of the radius ε containing less than minpts objects and which at the same time does not belong to the neighbourhood of any core object.

3.4.2.2 Optimization of the neighbourhood radius

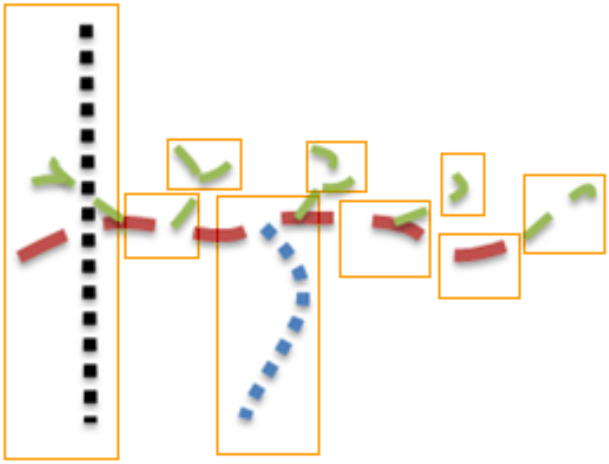
The neighbourhood radius ε is estimated for a data set with the same dimensionality

as the data studied, but uniformly distributed within the range of the experimental data (Daszykowski, Walczak & Massart 2001).

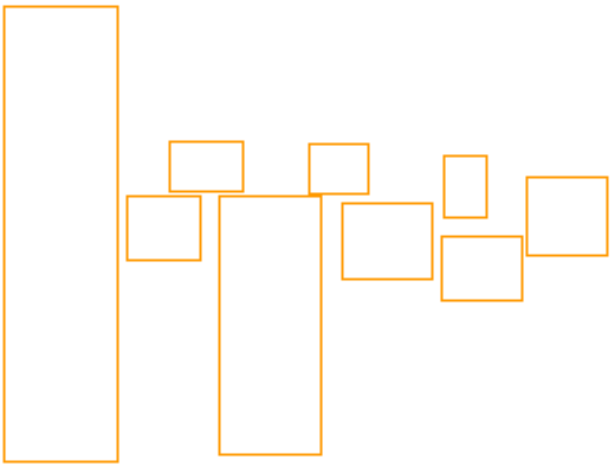
In order to choose the optimized ε for a data set $X\,(m \times m)$, a set of m objects is simulated in the n dimensional space in the range of X variables. Then the calculation of the distances between each object of its kth nearest neighbour is performed, where k is equal to minpts. The m calculated distances are sorted and then ε is selected as the distance equal to the 95% quantile (Quantiles are points taken at regular intervals from the cumulative distribution function (CDF) of a random variable) (Daszykowski, Walczak & Massart 2001).

3.4.2.3 Clusters Extracted

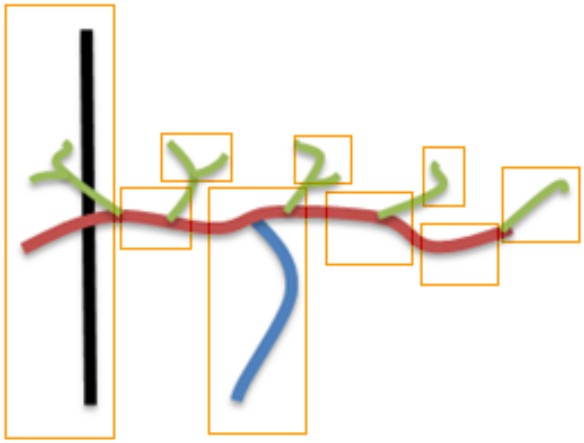
Clusters are extracted from the data obtained from cylinder feature extraction using the density clustering method of NP. The parameter minpts, the minimum number of points required to form a cluster, is set as 2 in order to minimize the noise points. Areas of each cluster can be calculated by the boundary points (Figure 3.8-a&b). Then, the original input points can be divided into several clusters through the areas (Figure 3.8-c).



(a) Clusters Obtained using NP



(b) Areas of Clusters



(c) Clusters of Original Input Points

Figure 3.8 Clusters Obtained via Density Clustering

Chapter 4 Object Identification and Location Derivation

After segmentation, the input data points are divided into clusters. Then the next goal is to identify the four key features of vineyards and derive their locations respectively. There are five steps of the proposed method including four features identification and location derivation as shown in the flow chart (Figure 4.1).

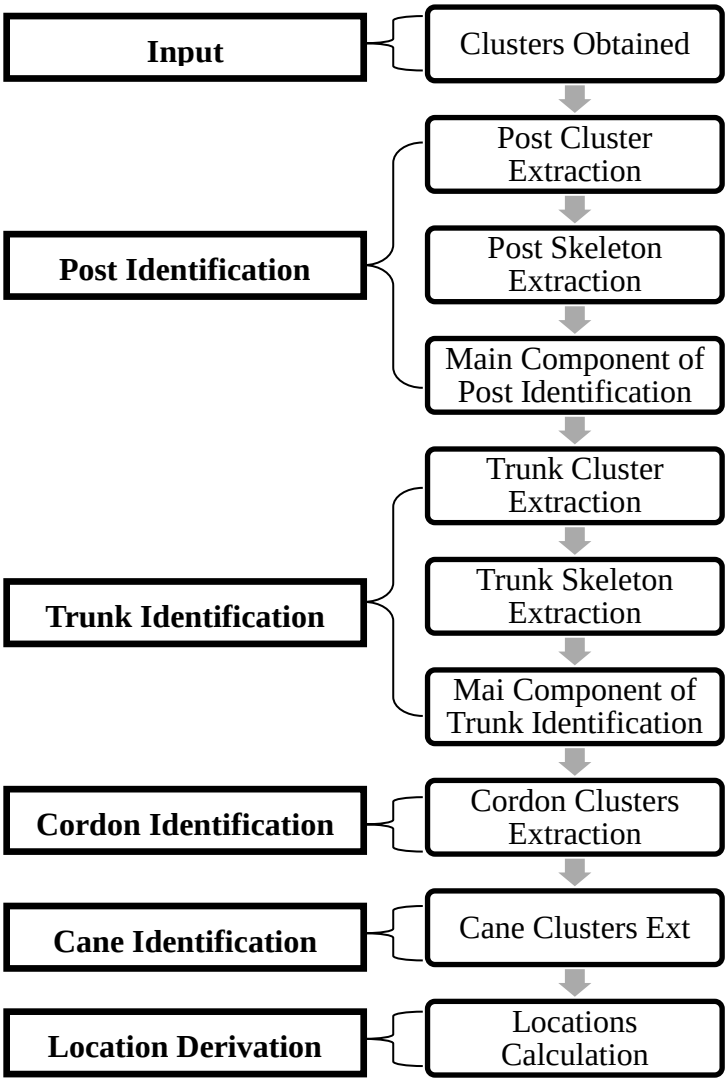


Figure 4.1 Flow Chart of Objects Identification and Location Derivation

The input, Clusters obtained, is the outcome of Segmentation (Chapter 3).

Step one is post identification (section 4.1). The post cluster is extracted by a point number threshold filtering. The cluster obtained contains not only the post points but also some other points. Therefore, a refinement process is performed including skeleton extraction and main component identification.

Step two is trunk identification (section 4.2) sharing the same sequences as step one.

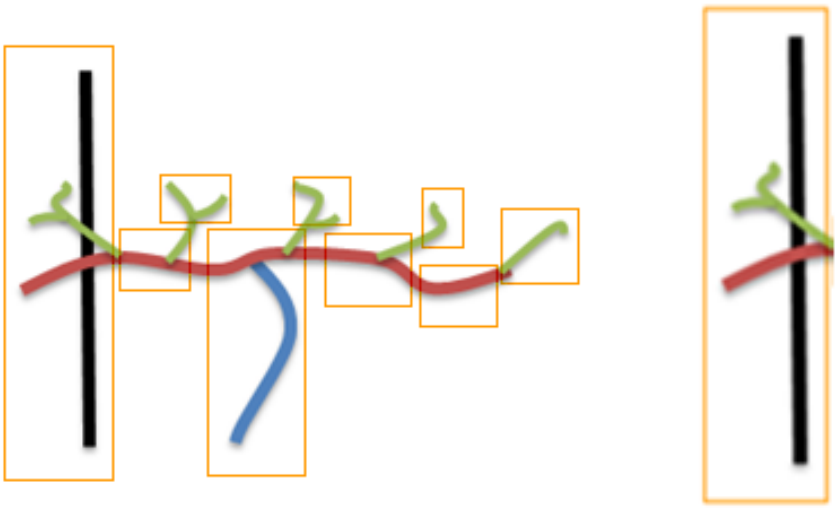
Step three is cordon identification (section 4.3). The extraction of cordon clusters is based on the trunk points identified in step two.

Step four is cane identification (section 4.4).

Step Five is location derivation (section 4.5) of all four features.

4.1 Post Identification

Each cluster obtained through segmentation contains a certain number of points. It is observed that posts are normally the highest and widest objects in the scene of vineyards, which means posts contain the highest number of points. By comparing with several input set points, a threshold of point number can be obtained. The clusters that represent the post area can be directly obtained via threshold filtering (Figure 4.2). However, the cluster representing a post not only contains points of the post but also other points including some parts of the cordon and cane which may affect the performance of the pruning operation. Therefore, a refinement process is necessary. To improve the accuracy of post identification, a two step method is proposed with skeleton extraction and main component identification.



(a) All Clusters Obtained (b) Post Cluster

Figure 4.2 Post Cluster Extraction

4.1.1 Skeleton Extraction

Point cloud skeleton extraction via Laplacian-based contraction (Cao et al. 2010) is applied here to extract the skeleton of the post, which has two steps including geometric contraction and topological thinning.

4.1.1.1 Geometric Contraction

The goal of the geometric contraction is to maintain the global shape of the input points through anchoring points chosen by an implicit Laplacian smoothing process (Cao et al. 2010). As a result, the geometric characteristics of the input points can be captured which helps estimate the position of the final curve skeleton. The key concept of geometric contraction is to solve the linear system described in equation (4.1):

$$\begin{bmatrix} W_L L \\ W_H \end{bmatrix} P' = \begin{bmatrix} 0 \\ W_H P \end{bmatrix} \qquad (4.1)$$

, where L is a $n \times n$ Laplacian matrix with cotangent weights; P is the cluster represent the post area; P' is a contracted point cloud; and W_L and W_H are the

diagonal weight matrices balancing the contraction and attraction constraints, the ith diagonal element of W_L (resp. W_H) is denoted $W_{L,i}$ (resp. $W_{H,i}$).

The solution to this system minimizes the quadratic energy (equation (4.2)):

$$\|W_L LP'\|^2 + \sum_i W_{H,i}^2 \|P_i' - P_i\| \qquad (4.2)$$

, where the first term removes geometry details along the normal directions using implicit Laplacian smoothing and the second preserves shape geometry during contractions.

In order to construct the Laplacian operator, an approximate neighbourhood of P_i is firstly extracted by computing k nearest neighbors $N_k(P_i)$ of P_i and projecting them on the tangent plane defined by their principal components. Then, a planar Delaunay triangulation is constructed and one-ring neighbours of P_i is defined. The neighborhood information is computed only in the beginning because it is difficult to compute a correct tangent plane as the contraction procedure goes. Therefore, subsequent iterations only need to update the Laplacian weights rather than the whole operator.

In order to collapse the post cluster P into a curve skeleton, the equation (4.2) is run iteratively while the two parameters W_H and W_L is updated each time. The contracted result P' is contracted notably after the first iteration and the contraction weights W_L is increased after each iteration. The attraction weights $W_{H,i}$ is also needed to be updated according to the collapsed degree of p_i in order to avoid over contraction. The purpose is accomplished by the extent of its neighbors that is approximated by $min_{q \in N_k(p_i)} \|p_i - q\|$.

In order to make points with smaller neighbourhoods be attracted more strongly to their current positions, the iteration t is evaluated. Firstly, the equation (4.3) is solved.

Then, $W_L^{t+1} = s_L W_L^t$ and $W_{H,i}^{t+1} = W_{H,i}^0 S_i^0 / S_i^t$, where S_i^t and S_i^0 are the current and original neighbourhood extent of point i. Finally, the new Laplacian operator L^{t+1} with the current point cloud P^{t+1} is computed (Cao et al. 2010).

$$\begin{bmatrix} W_L^t L^t \\ W_H^t \end{bmatrix} P^{t+1} = \begin{bmatrix} 0 \\ W_H^t P^t \end{bmatrix} \tag{4.3}$$

The outcome of this algorithm is a thin contracted point cloud C (P^{t+1}) that captures the geometric characteristics of the post cluster.

4.1.1.2 Topological Thinning

After the contracted point cloud C is obtained through the geometric contraction, the next step is to convert C to a curve skeleton using connectivity building and edge contraction.

In terms of connectivity building, C is sampled by using farthest point sampling and a ball of radius r. Each sample $g_i \in \mathrm{G}$ represents the set of associated points C_i in C which are closest to it, such that $\mathrm{C} = \bigcup_i C_i$ and $C_i \cap C_j = \emptyset,\ \forall i \neq j$. By connecting g_i and g_j if their associated points share common local 1 ring neighbours, a graph G is obtained with uniformly distributed nodes (Cao et al. 2010).

After the connectivity is built, unnecessary edges are collapsed until no triangles exist to build a curve skeleton. In order to distribute the final skeleton vertices uniformly and capture the shape of G, the edge with minimum Euclidean length to its midpoint is collapsed and triangles incident to the edge are removed. The points associated with the two endpoints are assigned to the newly created vertex. The procedure is iterated until all triangles are removed and the curve skeleton is built.

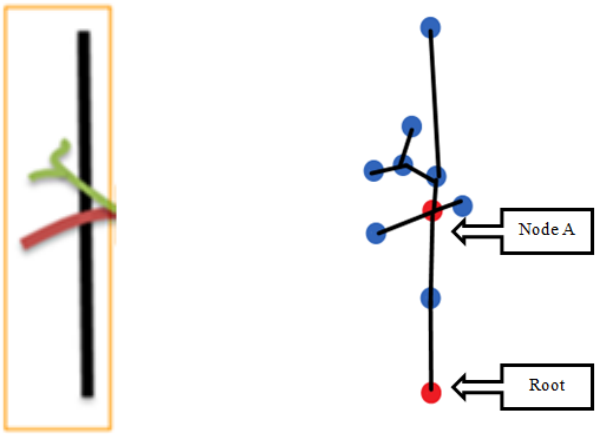
4.1.2 Main Component Identification

From the skeleton extracted, it is observed that the connection edges from root node

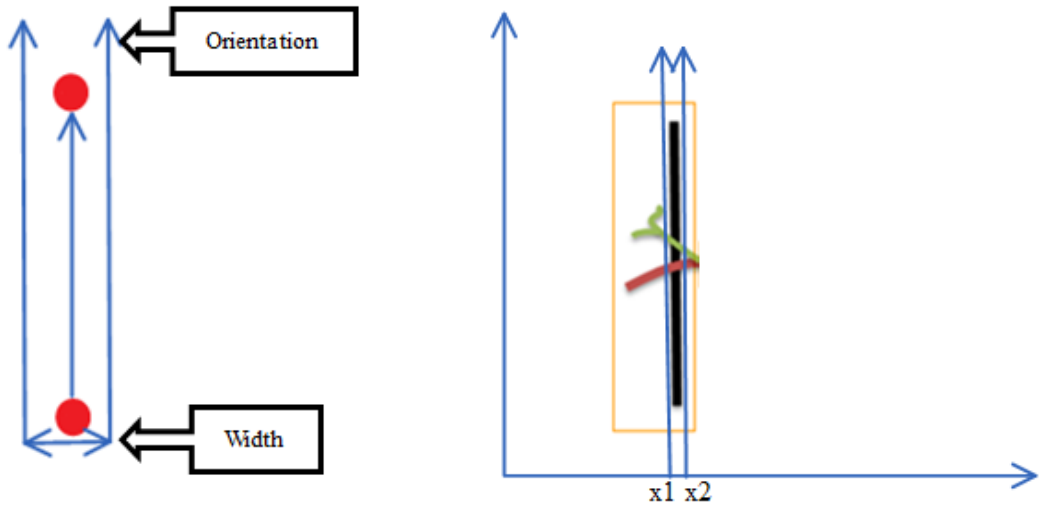
to node A which is the first node connects three nodes represent the key component of the post as shown in Figure 4.3-b. The main component represents the main part of the post and it contains most of the points belonging to the post. By analysing the main component, the whole post can be identified.

Firstly, the orientation and width of the post can be extracted by connecting the root and the node A (Figure 4.3-b). An axis can be built based on the orientation extracted. Then, all the points of the post cluster are rotated to the new axis, the vertical direction of which is the same as the orientation of the post. After that, the two values x1 and x2 as shown in Figure 4.3-d can be calculated by locating the boundary points of the main component. From this point, the threshold is built for identifying the whole post and a refinement operation can be performed. The points belonging to the post can be extracted and noisy points including some parts of the cordon and cane can be removed.

The orientation and the width of the post can be estimated through this key component. The points belonging to the post and noisy points can be identified by rotating the axis to the direction of the post. Then, points within the boundary from x1 to x2 belong to the post and those left are noise (Figure 4.3-d).



(a) Post Area with Noisy Points (b) Skeleton Extracted

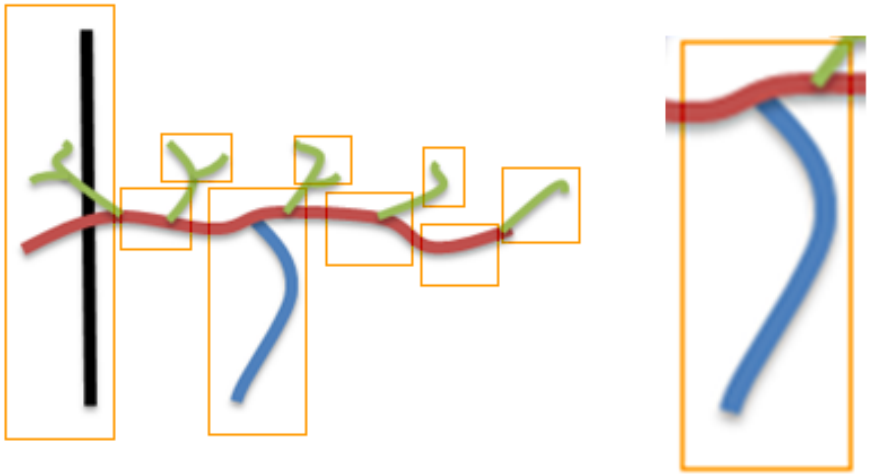


(c) Orientation and Width (d) Refinement of Post Identification

Figure 4.3 Post Identification and Refinement

4.2 Trunk Identification

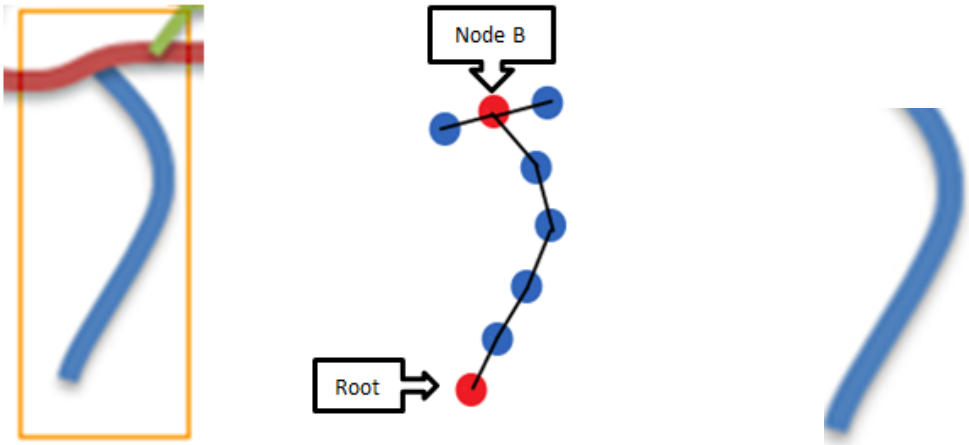
Each cluster obtained through segmentation contains a certain number of points. It is observed that trunks are the second highest and widest objects in the scene, which means trunks contain the second highest number of points. By comparing with several input set points, a threshold of point number can be obtained. The clusters that represent the trunk area can be directly obtained via threshold filtering (Figure 4.4). However, the trunk cluster sometimes contains some parts of the cordon and cane and a refinement operation is required.



(a) Clusters Obtained (b) Trunk Cluster

Figure 4.4 Trunk Cluster Extraction

To refine the trunk area and improve the accuracy of the trunk identification, the skeleton extraction method of the post is applied again. After applying geometric contraction and topological thinning, the skeleton of the trunk cluster is extracted as shown in Figure 4.5. It is observed that the nodes connection from root node to node B, which is the first node connecting with 3 nodes, represents the trunk. As a result, the trunk is identified and refined and the noisy points are removed.

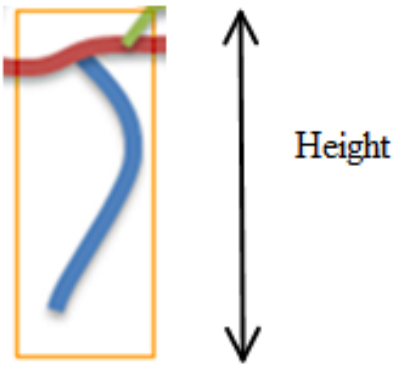


a. Trunk Cluster b. Skeleton of Trunk c. Trunk Refinement

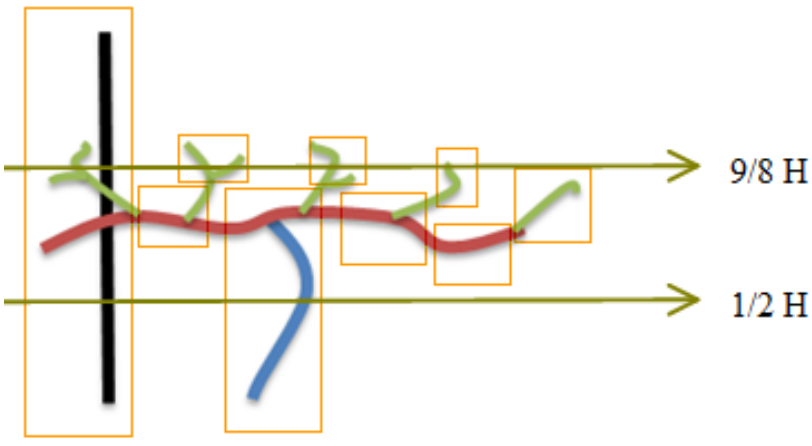
Figure 4.5 Trunk Identification and Refinement

4.3 Cordon Identification

Based on the natural feature of grapevines, cordon area threshold is obtained by measuring and analysing the input data sets. It is observed that the cordon area is between 1/2 and 9/8 of the total height of the trunk. Therefore, all the clusters within this area are identified as points of cordons (Figure 4.6). To improve the accuracy of this progress, the cluster which has 50 per cent of points within this area is also identified as cordon.



(a) Height of Trunk



(b) Cordon Area

Figure 4.6 Cordon Area Estimation

4.4 Cane Identification

As features of post, trunk and cordon are extracted and points belonging to them are identified, the points left contain all canes. Therefore, cane identification is automatically achieved.

4.5 LOCATION DERIVATION

The input data used in this thesis is point clouds which is a set of vertices in a three dimensional coordinate system. These vertices contain the location information of the four features in vineyards. As the points (vertices) belonging to post, trunk, cordon and cane are identified, the locations of them can be derived directly. The location of post can be derived by calculating a cuboid defined by the boundary points of left, right, top, bottom and back as well as the locations of trunk, cordon and cane. This is one benefit of using point clouds as input rather than images obtained by machine vision, the only focus is how to extract features and the calculation of location is not needed.

Chapter 5 Experimental Work

In order to test and analyse the effectiveness and performance of the proposed method, experiments are conducted. The experimental setup is introduced in section 5.1 including the range sensor applied and the program used to capture point clouds. In section 5.2, the experiment results are illustrated and analysed using several groups of data sets captured in different vineyard scenes.

5.1. Experimental Setup

5.1.1 Kinect

Kinect (Figure 5.1) is a motion sensing input device developed by Microsoft for Xbox 360, video game consoles and Windows PCs. It enables users and researchers to control and interact with consoles and computers as a game controller. The application of the Kinect is based on a webcam style add-on peripheral and a natural user interface using gestures and spoken commands. One feature of the Kinect is the ability to capture point cloud data with a fast and accurate performance. In this thesis, the Kinect is used as the input data capturing range sensor to perceive the environment of vineyards.

Figure 5.1 The Kinect

The practical ranging limit of the Kinect sensor is from 1.2 to 3.5 meters with the Xbox software and Mircosoft SDK for Windows. It has the angular field of view of 57

degrees horizontally and 43 degrees vertically. The motorized pivot is capable of tilting the sensor up or down 27 degrees.

5.1.2 Data Capturing Program

Microsoft released a non-commercial Kinect software development kit (SDK) for Windows 7 in 2011, which includes compatible PC drivers for Kinect devices in Windows 7 environment. The SDK provides Kinect capabilities to developers to write and build applications and programs using C++, C# and Visual Basic programing languages through Microsoft Visual Studio 2010. The SDK includes the features:

- Raw sensor streams: Access to low-level streams from the depth sensor, color camera sensor, and four-element microphone array.

- Skeletal tracking: The capability of tracking the skeleton image of one or two people moving within Kinect's field of view for gesture-driven applications.

- Advanced audio capabilities: Audio processing capabilities include sophisticated acoustic noise suppression and echo cancellation, beam formation to identify the current sound source, and integration with Windows speech recognition API.

Under the consideration of compatibility and engineering implementation, the data capturing program is written in C # programing language using the Kinect SDK of 1.0 version. The interface of the program is illustrated in Figure 5.2 (initializing mode) and Figure 5.3 (capturing mode). The program is capable of capturing 2D colour images, 2D images with depth information and point clouds.

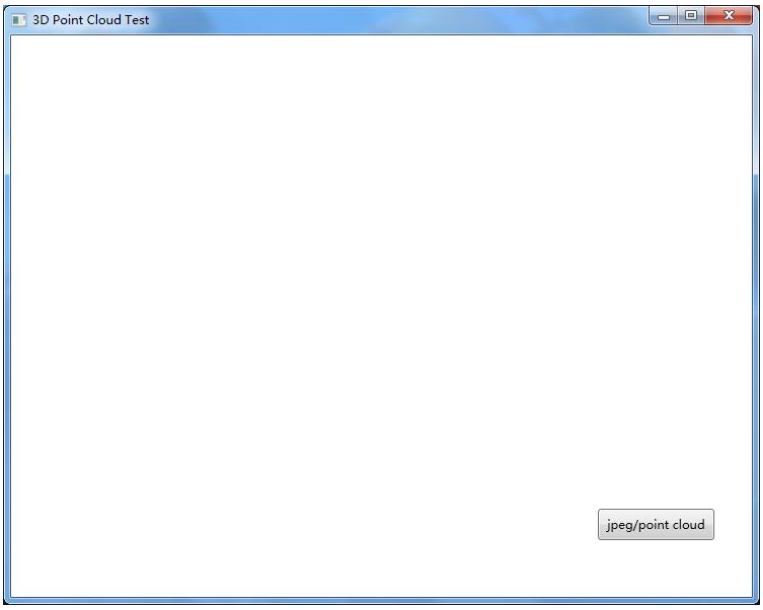


Figure 5.2 Data Capturing Program User Interface – Initializing Mode

Figure 5.3 Data Capturing Program User Interface – Capturing Mode

The program is used on a Windows 7 laptop of Thinkpad X201. The laptop has an Intel(R) Core(TM) i5 CPU M520 with 2.40GHz, a 4G RAM and Intel(R) HD Graphics 4000.

5.1.3 Field Configurations

The point clouds used in this paper are captured in the vineyard at the Waite Campus,

the University of Adelaide. The field of view is illustrated in Figure 5.3. In the experiment, the Kinect sensor is located at the height of 1.5 meters above the ground and the distance of 1 meter to the grapevine. The Kinect sensor is placed horizontally with an angular field of view of 57° horizontally and 43° vertically.

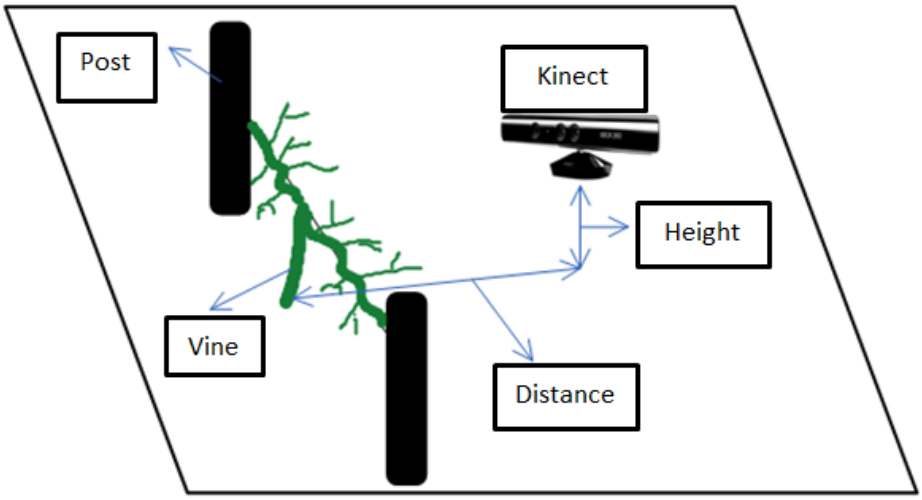


Figure 5.4 Field View

5.1.4 Method Testing Program

The proposed method is written in Matlab code in order to test and analyse its effectiveness and performance with the captured input point clouds. The 6.5 version of the Matlab software is selected and running in the Windows XP environment. The Dell OptiPlex 780 runs the Matlab code has the Intel(R) Core(TM)2 Quad CPU Q9400 with 2.66GHz, a 4G RAM and the Graphic card of ATI Radeon HD 3450.

5.2 Experimental Results

The method proposed was applied to 20 data sets captured in the vineyard aforementioned. Six data sets, which represent six situations of input points in vineyards, are selected to illustrate the effectiveness of the proposed method step by step. The accuracy and performance of the method are illustrated as well. All the data sets are tested and analysed following the sequences of the proposed method.

5.2.1 Data Set One

5.2.1.1 Data Set One Input

Data set one represents one of the most common scenes obtained from input data sets. As illustrated in Figure 5.5, there are one post, one trunk, several cordons unconnected and scattered canes captured. Besides, the post and trunk are very close to and overlapped with each other which make the scene very complex.

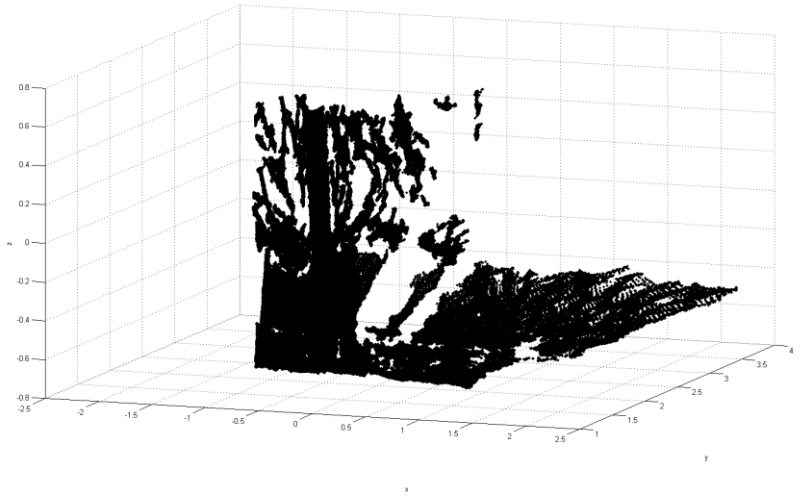
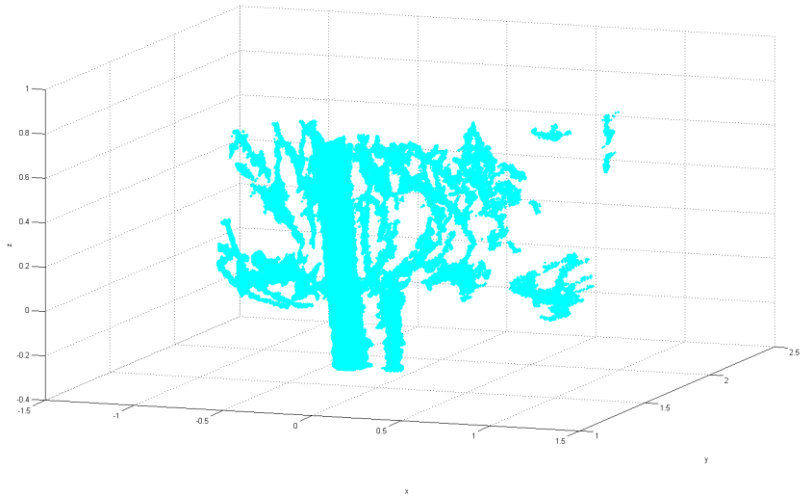


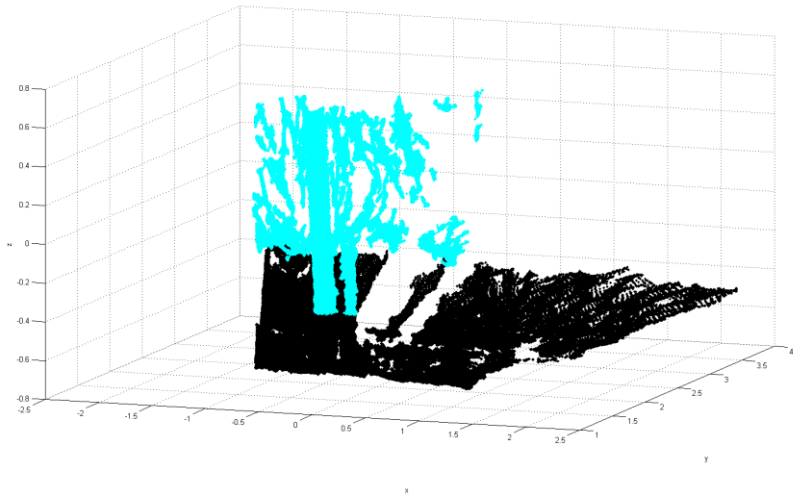
Figure 5.5 Original Input of Data Set One

5.2.1.2 Data Set One Segmentation

The segmentation is performed when the input is obtained. Firstly, pre-processing is applied in order to remove unwanted points including the solid ground and other rows of grapevines. In the case of data set one, only solid ground is the noisy data. The pre-processing result is shown in Figure 5.6.



a. Pre-processing of Data Set One



b. Pre-processing of Data Set One in the Original Input

Figure 5.6 Data Set One Pre-processing

Then, the normal vectors of the pre-processed input are estimated using KNN and by using of which the Gaussian sphere is generated as shown in Figure 5.7. In order to generate the Hough Gaussian sphere, the Gaussian sphere is transformed via the Hough transform. As shown in Figure 5.8, there are many great circles on the unit sphere and they are intersected with each other. One of the intersections represents the orientation of the post.

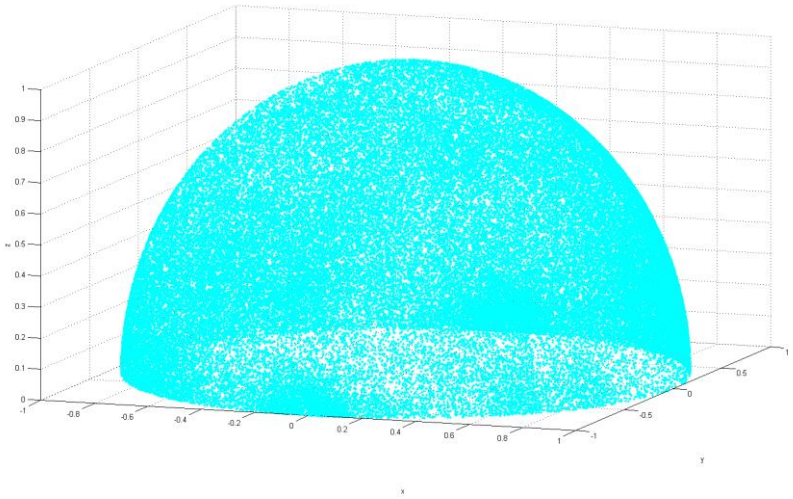


Figure 5.7 Gaussian Sphere of Data Set One

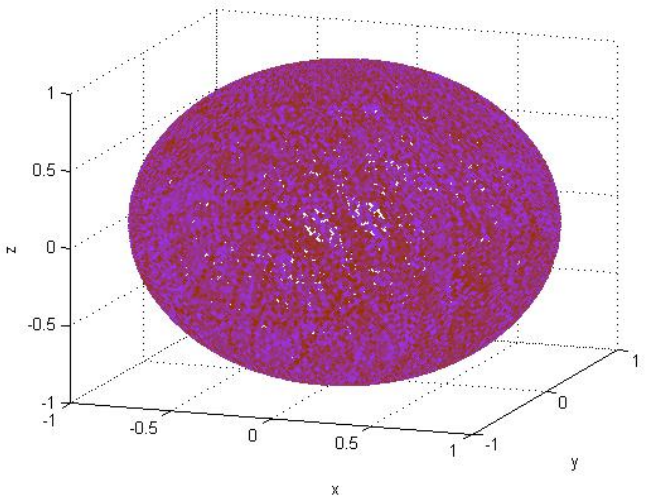


Figure 5.8 Hough Gaussian Sphere of Data Set One

In order to estimate the orientation of the post, accumulation is performed to locate the intersection that represents the post. Great circles vote for the cell that has the number of accumulator values that represent the post can be found. By using the great circles located, the cylinder feature is extracted. The points representing the possible points of the post can be extracted. Because some of the points belonging to the trunk, cordon and cane also vote for the accumulator value of the post, they are also extracted. As shown in Figure 5.9, more points belonging to the post and trunk are extracted than the points of cordons and canes, which make the density feature of post

and trunk notable.

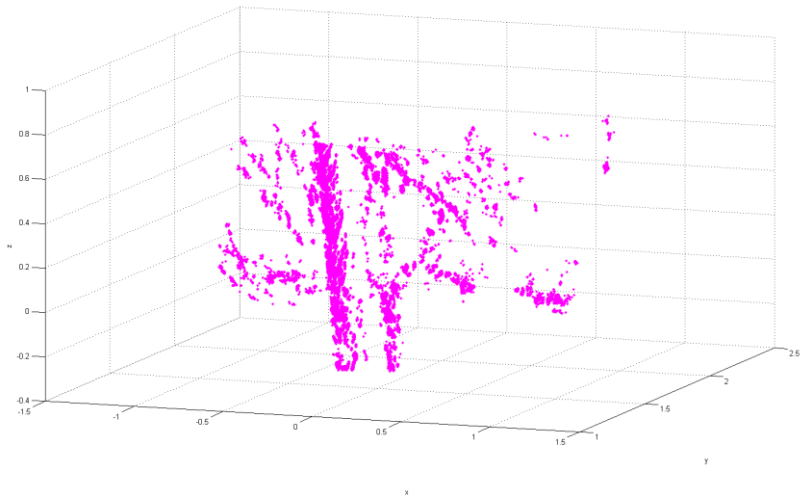


Figure 5.9 Cylinder Feature Extraction of Data Set One

As the density feature is notable enough after cylinder feature extraction, the density clustering operation is performed in order to divide the points into several clusters. As shown in Figure 5.10, the clusters obtained are demonstrated in the cubic form for better understanding and view. All the cubes (Figure 5.11) extracted are applied to the pre-processed input in order to obtain the clusters of the input as shown in Figure 5.12.

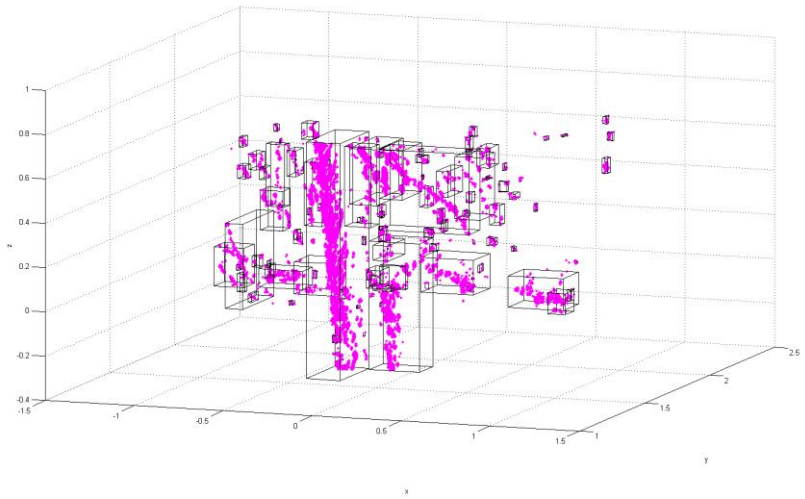


Figure 5.10 Density Clustering of Data Set One

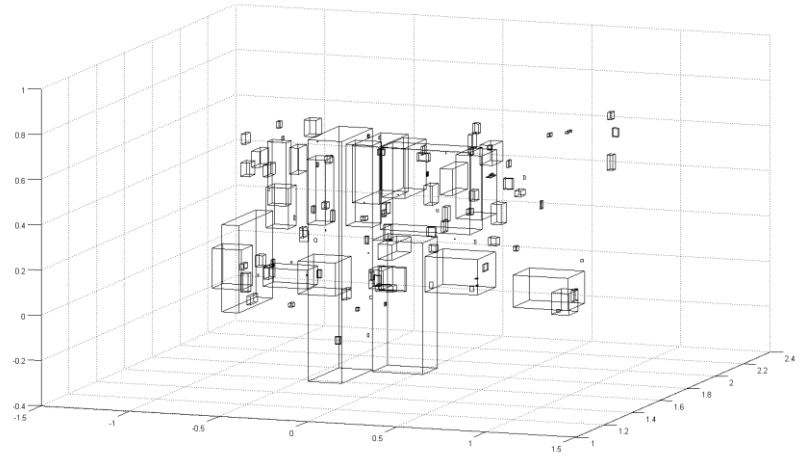


Figure 5.11 Cubic Extraction of Data Set One

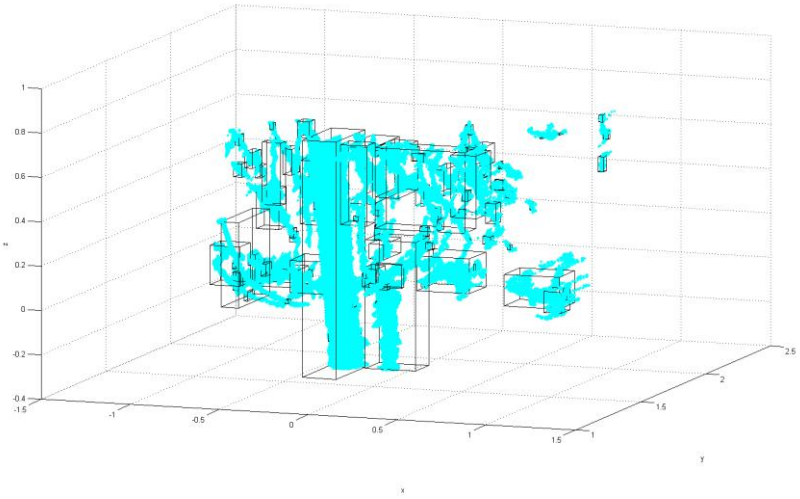


Figure 5.12 Clusters of the Pre-processed Input Points of Data Set One

5.2.1.3 Data Set One Object Identification

For each cluster obtained through segmentation contains a certain number of points. By comparing with several input set points, a threshold of point number can be obtained which is the point number larger than 18000. The cluster that contains this certain number of points can be directly obtained through threshold filtering (Figure 5.13).

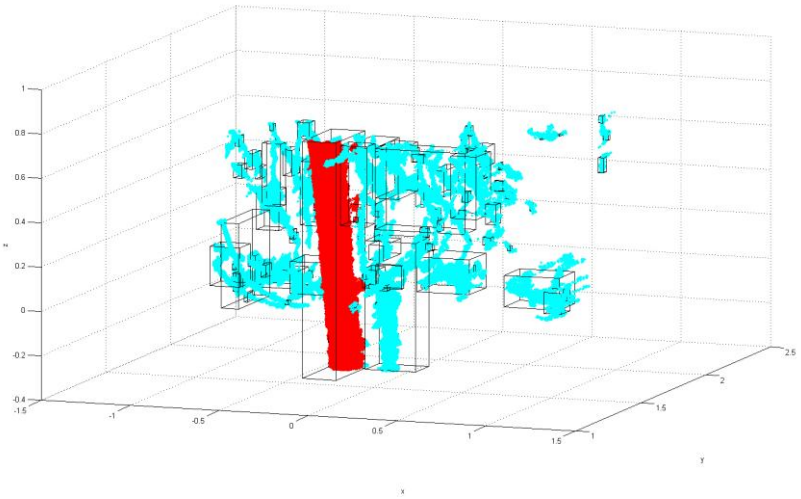


Figure 5.13 The Post Cluster (red colour) of Data Set One

In Figure 5.14, the close view of the post cluster is shown. As there is only a small number of noisy points in the post cluster, in order to determine whether a refinement operation is needed, the skeleton of the post cluster is extracted as shown in Figure 5.15. According to the proposed method, there is no joint that connects to more than two joints, which means the refinement operation is not necessary.

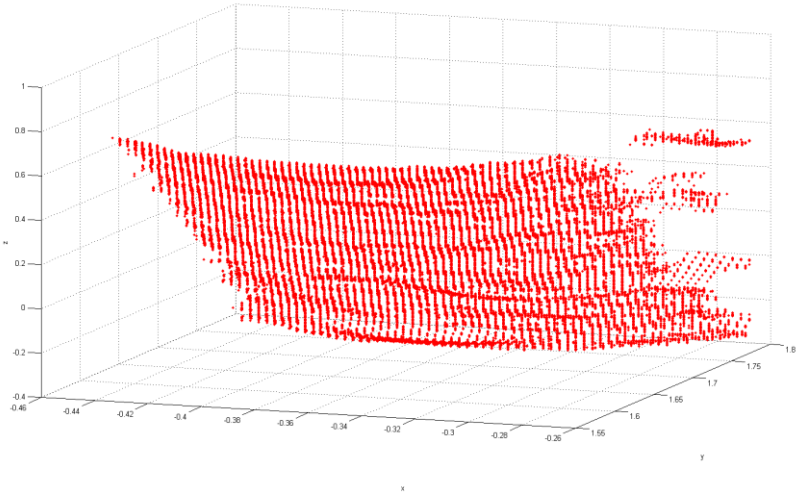


Figure 5.14 Close View of the Post Cluster of Data Set One

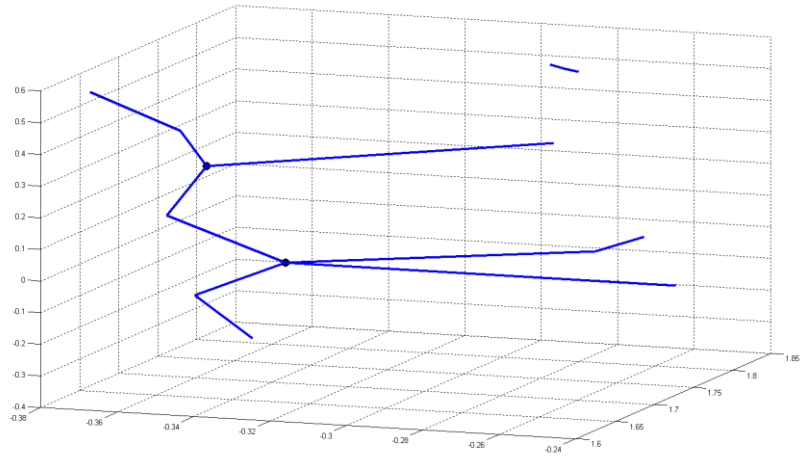


Figure 5.15 Skeleton Extraction of the Post Cluster of Data Set One

After the post is extracted, the next step is to extract the trunk. By using the same threshold filtering method, the clusters contain the number of points from 6000 to 13000 represent the trunk clusters. As there is only one trunk in the scene of data set one, the only trunk cluster is extracted as shown in Figure 5.16.

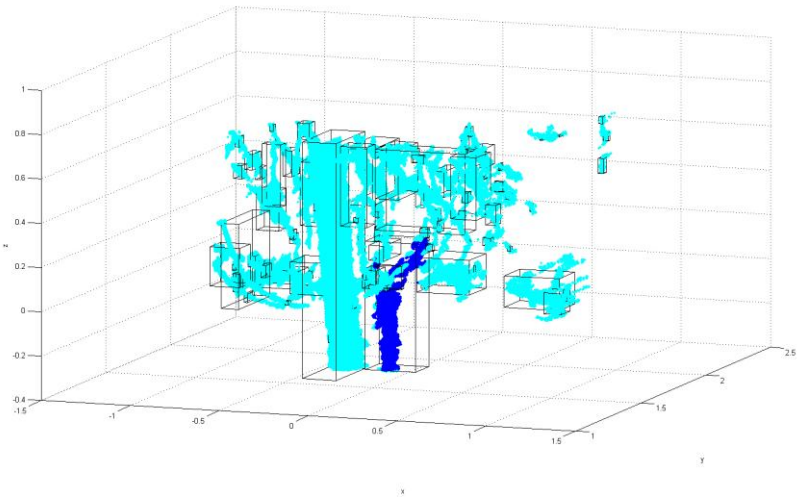


Figure 5.16 The Trunk Cluster of Data Set One (blue colour)

In Figure 5.17, the close view of the trunk cluster is shown. Except for the points belonging to the post, there are also some noisy points. For refinement purposes, the skeleton of the trunk is extracted as shown in Figure 5.18. The same as the post refinement, there is also no joint that connects to more than two joints, which means

no refinement operation can be performed. Therefore, the trunk is directly extracted.

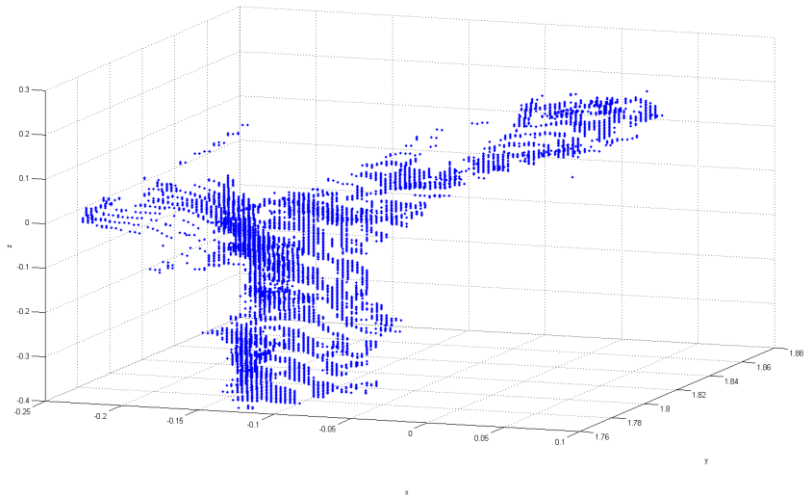


Figure 5.17 Close View of the Trunk Cluster of Data Set One

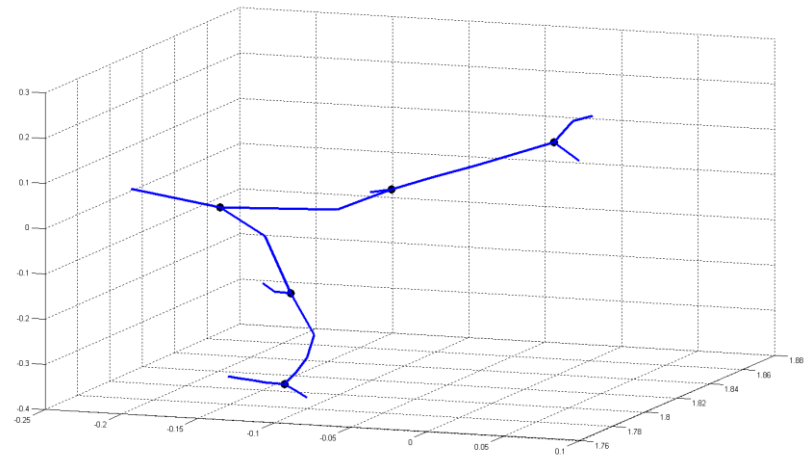


Figure 5.18 Skeleton of the Trunk of Data Set One

After trunk extraction, the cordon identification is performed. Based on the observation of all the data sets obtained and the natural feature of grapevines, the cordon area is within the vertical area between 1/2 and 9/8 of the total height of the trunk. As a result, all the clusters within this area are extracted. Besides, the clusters having 50 per cent of points within this area is also identified as the cordon to improve the accuracy of this process. Therefore, the cordon is extracted by combining all the clusters as shown in Figure 5.19. The accuracy of the trunk identification can

be improved by removing the points of the trunk within the area of the cordon.

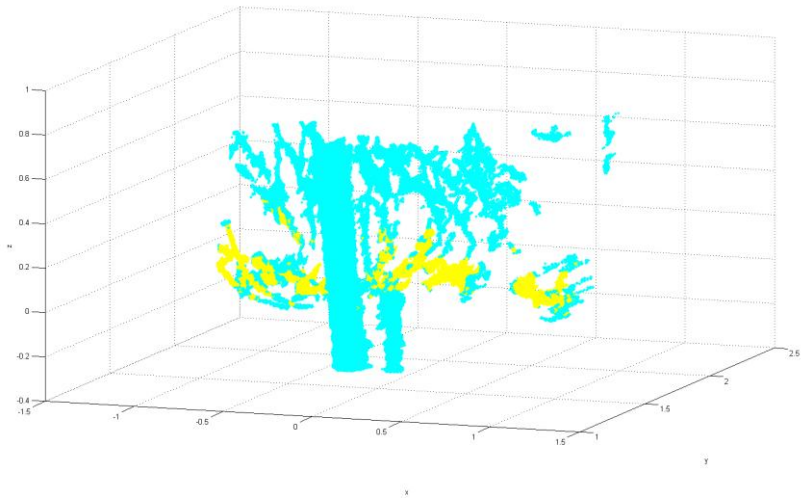


Figure 5.19 Cordon Extraction of Data Set One (yellow colour)

As there are only four objects in the pre-processed input, which are post, trunk, cordon and cane. Therefore, the left clusters are the points belonging to the cane (Figure 5.20).

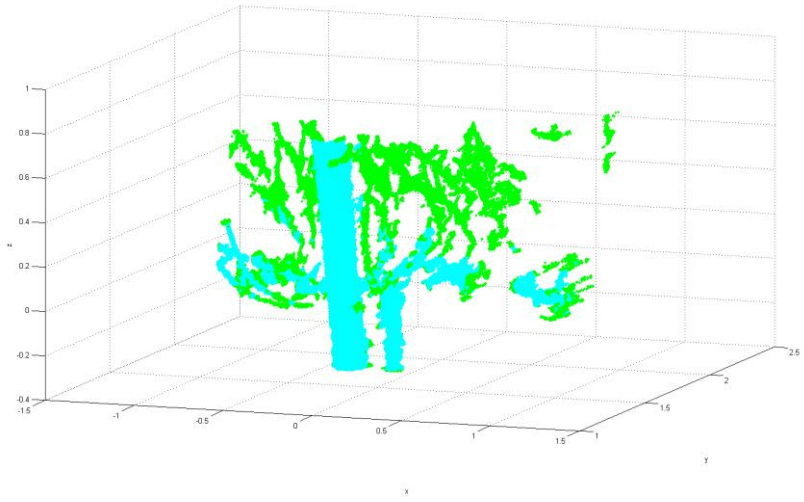
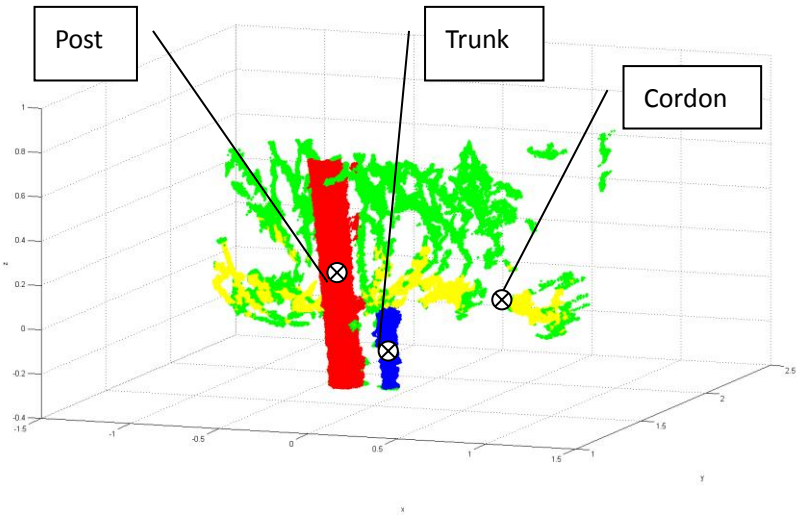


Figure 5.20 Cane Extraction of Data Set One (green colour)

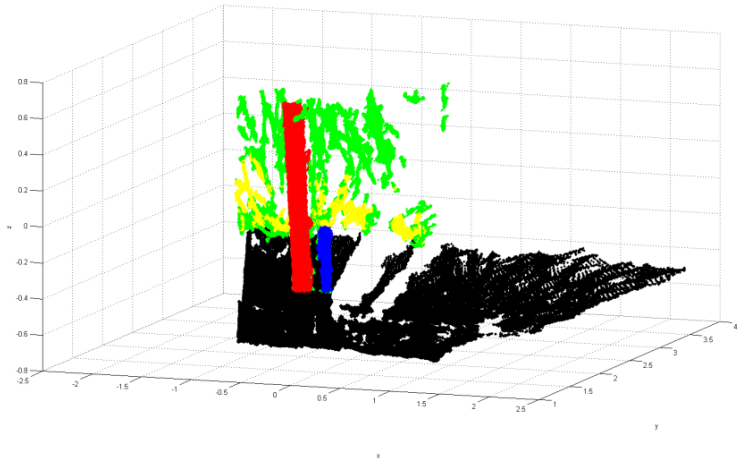
5.2.1.4 Data Set One Final Result

The final result of the proposed method applied on data set one is illustrated in Figure 5.21. The post, trunk, cordon and cane are identified as well as points belonging to

them. The positions of post and trunk are estimated by calculating their centres using the coordinates of the points extracted. The cordon is pruned part by part during vineyard pruning, position of one part of the cordon is estimated by calculating the centre of the selected part as shown in Figure 5.21-a. The unit of all the positions is meters. The position of post is (-0.3536, 1.6957, 0.0725). The position of trunk is (-0.1341, 1.7831,-0.4139). The selected part of cordon is from 0 of x axis to 1 of x axis and the position is (0.4089, 1.8343, 0.0146).



a. Object Identification of Data Set One in the Pre-processed Input

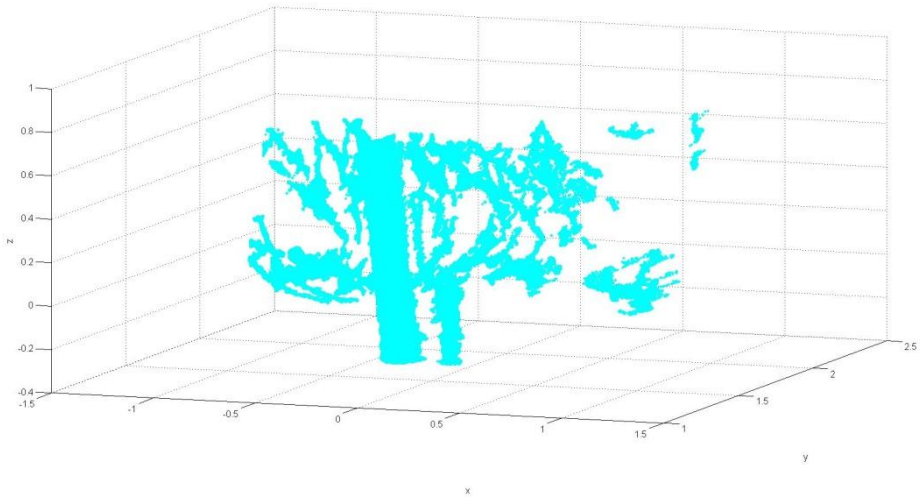


b. Object Identification of Data Set One in the Original Input

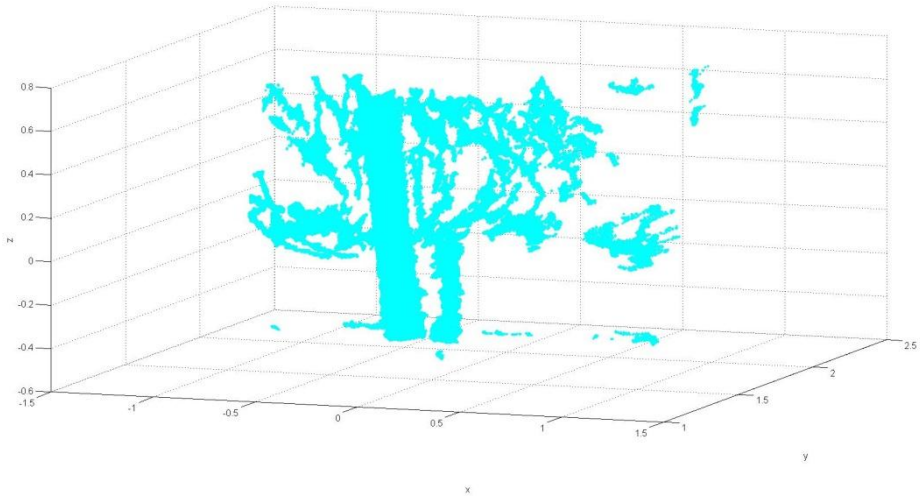
Figure 5.21 Final Result of Data Set One

5.2.1.5 Sensitivity Analysis of Thresholding in Pre-processing

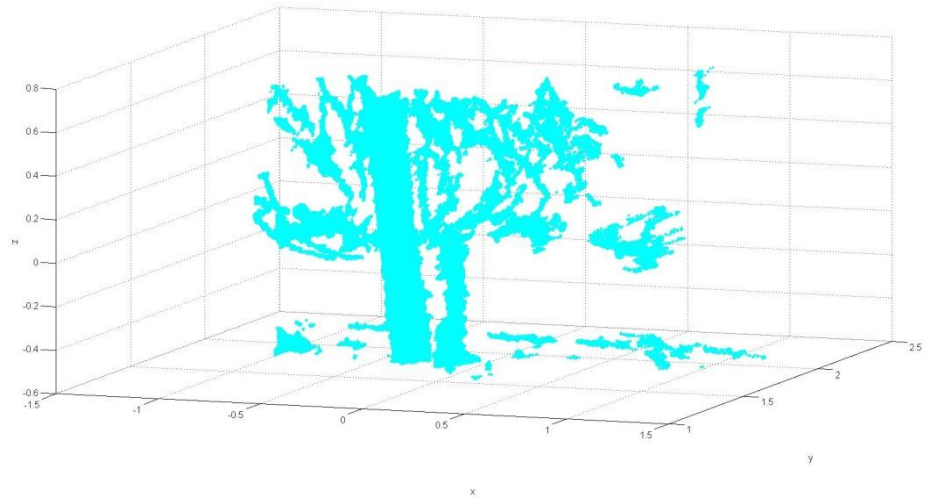
In order to analyse the sensitivity of thresholding method used in pre-processing, three different pre-processing results using the same original input are extracted via different threshold values. The main noisy points need to be removed are belonging to the ground; therefore, three values of Z axis are selected to perform the thresholding filtering where all the points below the threshold values along the Z axis are removed. The results are shown in Figure 5.22, where threshold value one equals 0.4, threshold value two equals 0.5 and threshold value three equals 0.6.



a. Threshold Value One (0.4)



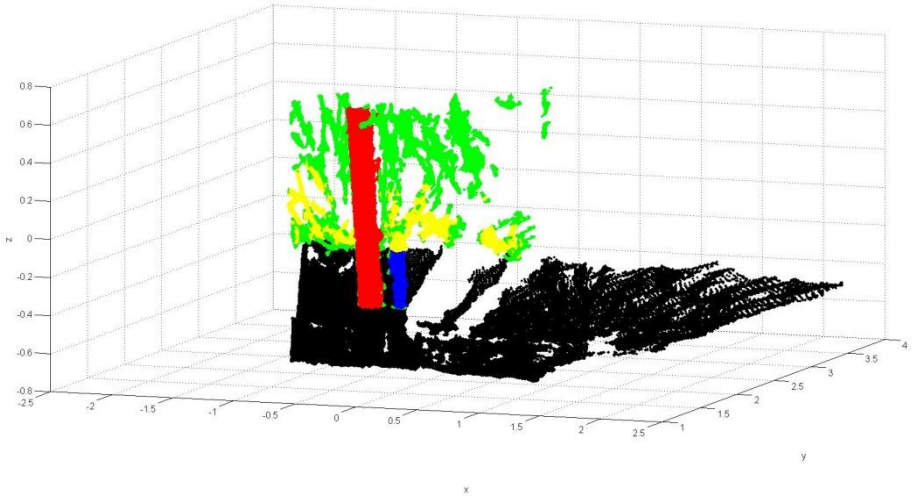
b. Threshold Value Two (0.5)



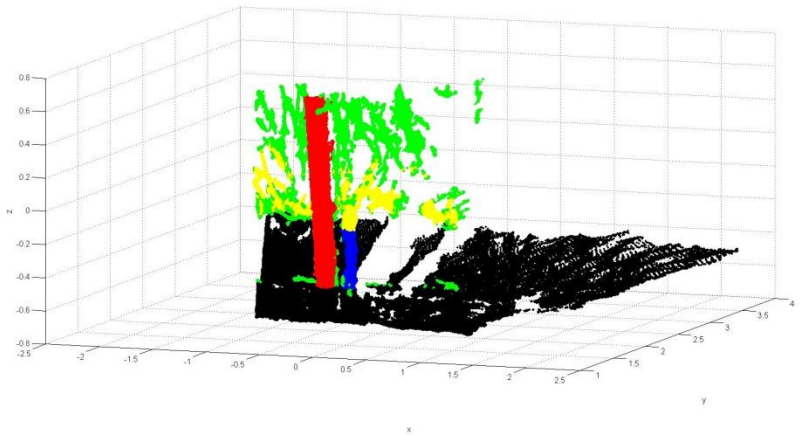
c. Threshold Value Three (0.6)

Figure 5.22 Results with Different Threshold Values

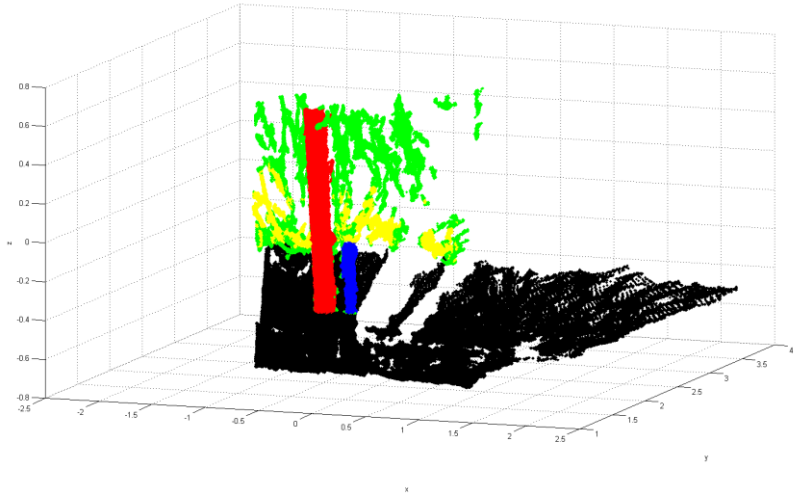
After processing, as shown in Figure 5.23, post, trunk, cordon and cane can all be correctly identified and extracted for all three sets of data, which demonstrates the robustness of the method. As the distance between two rows of grapevines is about 1.5m, the distance threshold value is not expected to be sensitive as long as it is set to be the distance between two rows of grapevines as demonstrated above.



a. Result of Threshold Value One (0.4)



b. Result of Threshold Value Two (0.5)



c. Result of Threshold Value Three (0.6)

Figure 5.23 Final Results Extracted for Different Thresholidng Outcomes

5.2.2 Data Set Two

5.2.2.1 Data Set Two Input

Data set two is another one of the most common scenes obtained from input data sets. The difference between it and data set one is the post and trunk are not close to and intersect with each other. It is also consist of one post, one trunk, several cordons

unconnected and scattered canes captured as shown in Figure 5.24.

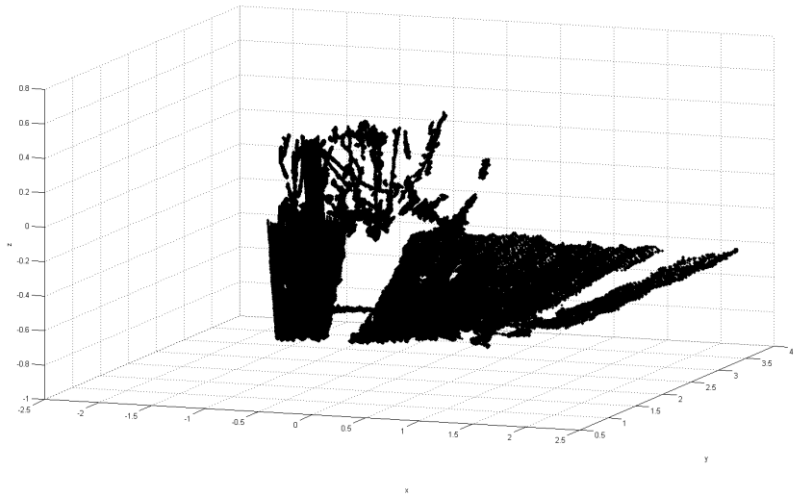
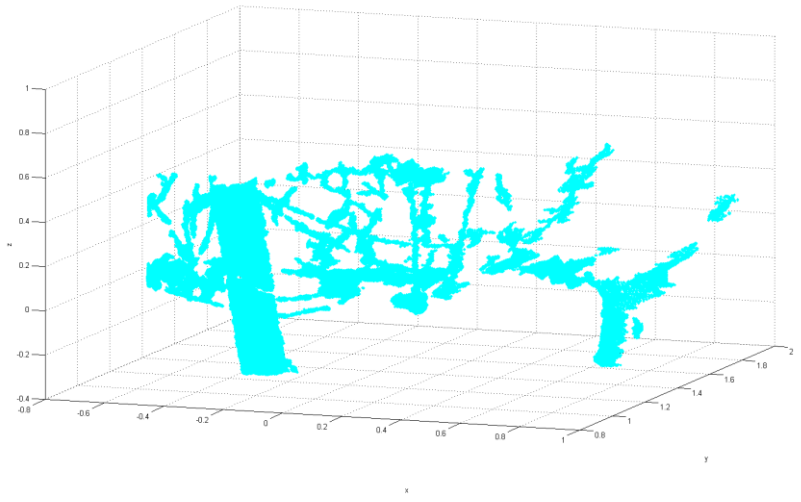


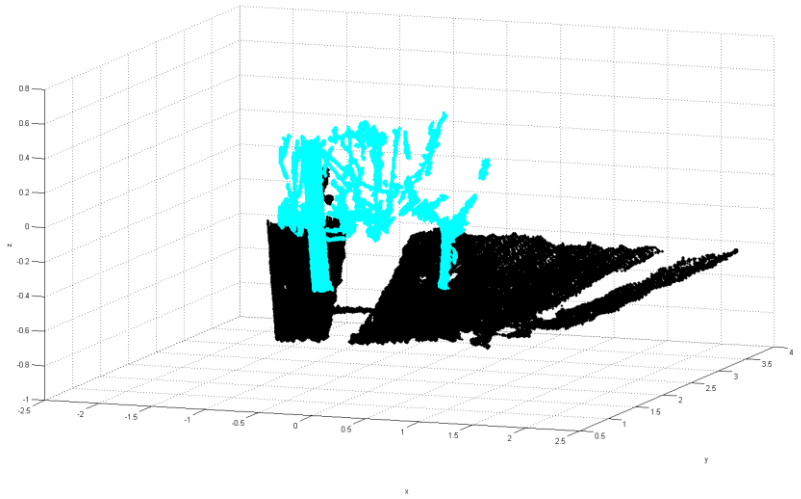
Figure 5.24 Original Input of Data Set Two

5.2.2.2 Data Set Two Segmentation

Firstly, pre-processing is performed and the result is shown in Figure 5.25.



a. Pre-processing of Data Set Two



b. Pre-processed Outcome in the Original Input of Data Set Two

Figure 5.25 Data Set One Pre-processing

Then, the normal vectors are estimated and the Gaussian sphere is generated as shown in Figure 5.26. The Hough Gaussian sphere (Figure 5.27) is also calculated based the result of Gaussian sphere.

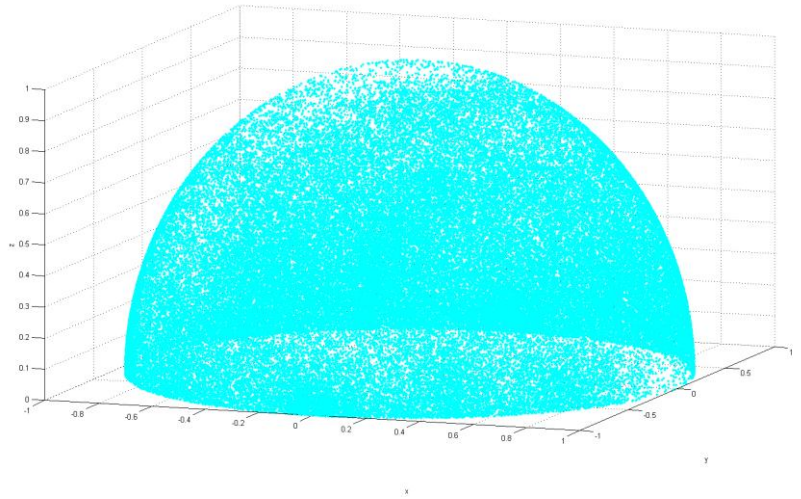


Figure 5.26 Gaussian Sphere of Data Set Two

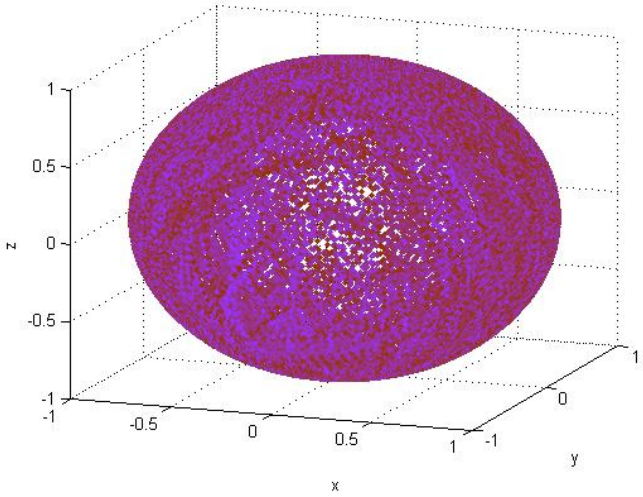


Figure 5.27 Hough Gaussian Sphere

After that, the cylinder extraction is performed and the result is shown in Figure 5.28.

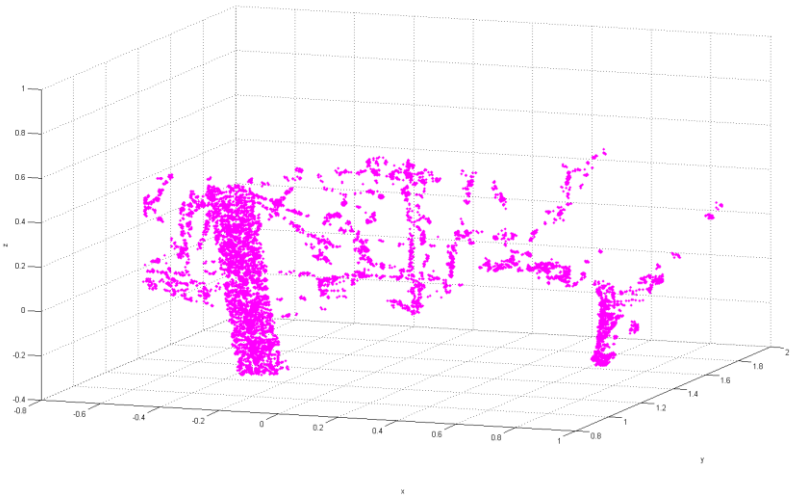


Figure 5.28 Cylinder Feature Extraction of Data Set Two

As the density feature is notable enough after cylinder feature extraction, the density clustering operation is performed in order to divide the points into several clusters. As shown in Figure 5.29, the clusters obtained are demonstrated in the cubic form for better understanding and view. All the cubes (Figure 5.30) extracted are applied to the pre-processed input in order to obtain the clusters of the input as shown in Figure 5.31.

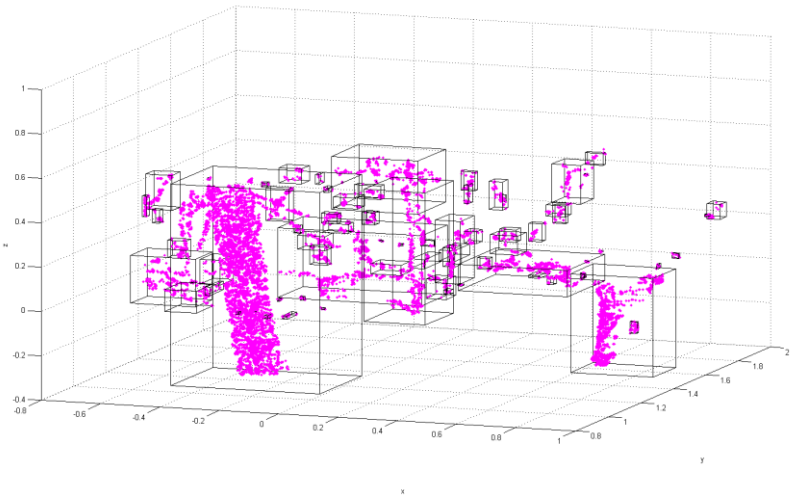


Figure 5.29 Density Clustering

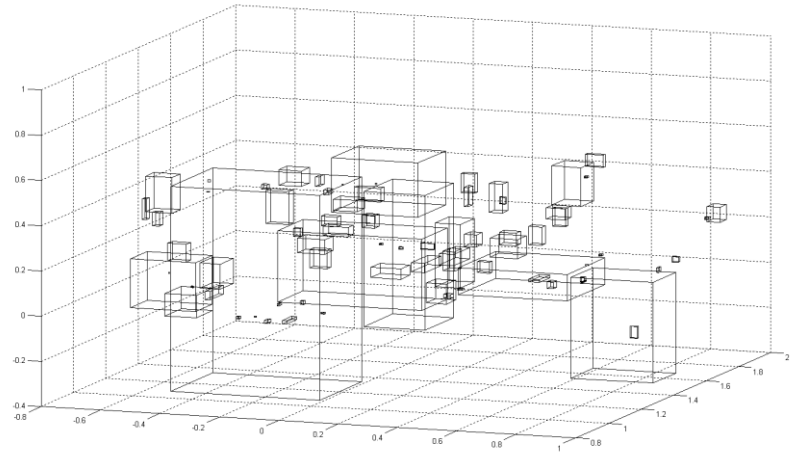


Figure 5.30 Cubic Extraction

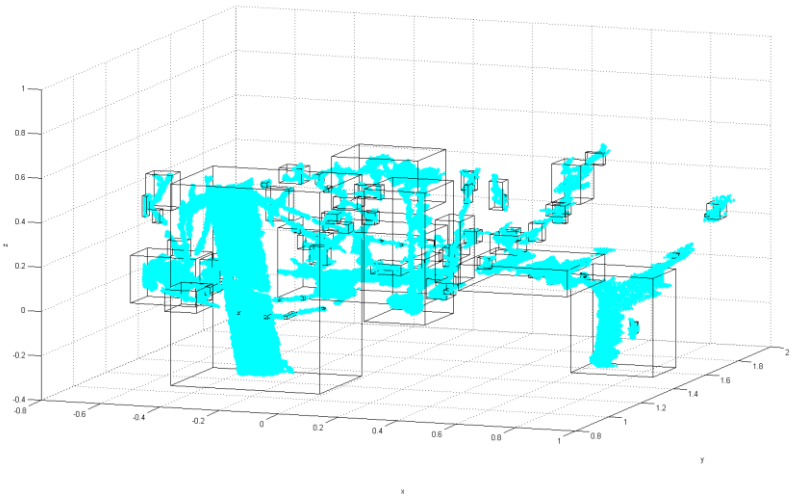


Figure 5.31 Clusters of the Pre-processed Input Points

5.2.2.3 Data Set Two Objects Identification

Firstly, the post cluster is extracted from the clusters obtained before as shown in Figure 5.32. As illustrated in Figure 5.33 of the close view of the post cluster, there are some noisy points need to be removed. In order to improve the accuracy of the post identification, skeleton of the post is extracted as shown in Figure 5.34 which shows the difference between the situation of Data Set Two and Data set one. There are several joints which connect with more than two joints. According to the proposed method, the main component extraction is performed. It is the part from the root joint to the first joint that connects with more than two joints. The key joint is extracted and shown in Figure 5.35. By locating the connecting relationship between the root joint and the key joint, the main component of the post cluster is extracted and shown in Figure 5.36. Then, a new axis is built by rotating the original axis to the direction defined by the line between the root joint and the key joint as shown in Figure 5.37. A threshold can be calculated by computing the width of the main component under the new axis. By threshold filtering (Figure 5.38), the noisy points can be removed and the accuracy of post identification is improved. The post identification result is shown in Figure 5.39.

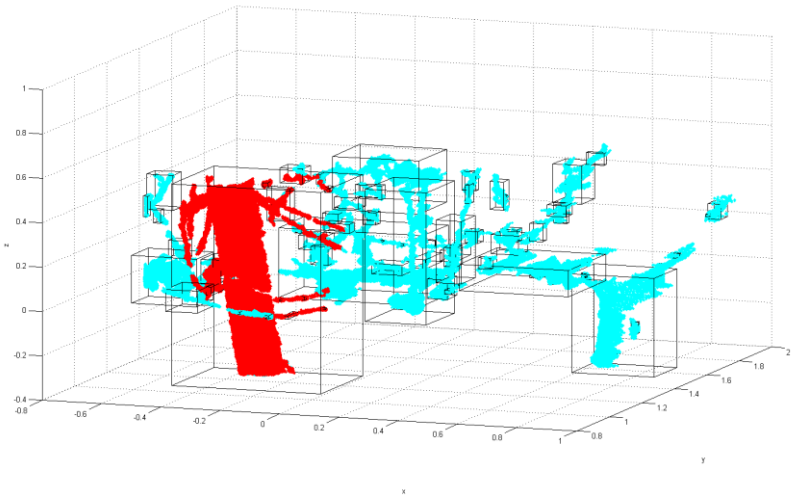


Figure 5.32 The Post Cluster of Data Set Two (red colour)

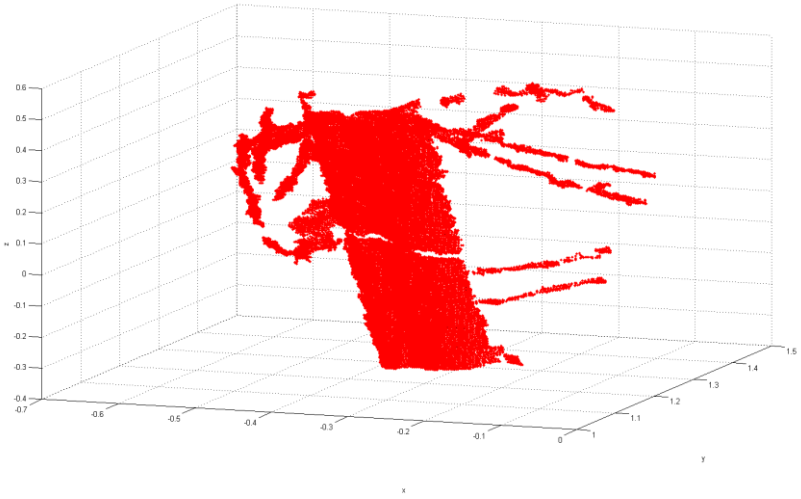


Figure 5.33 Close View of the Post Cluster of Data Set Two

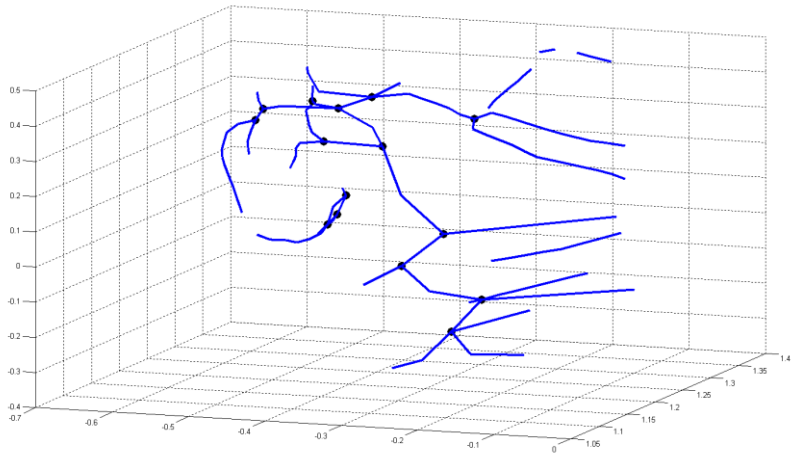


Figure 5.34 Skeleton Extraction of the Post Cluster of Data Set Two

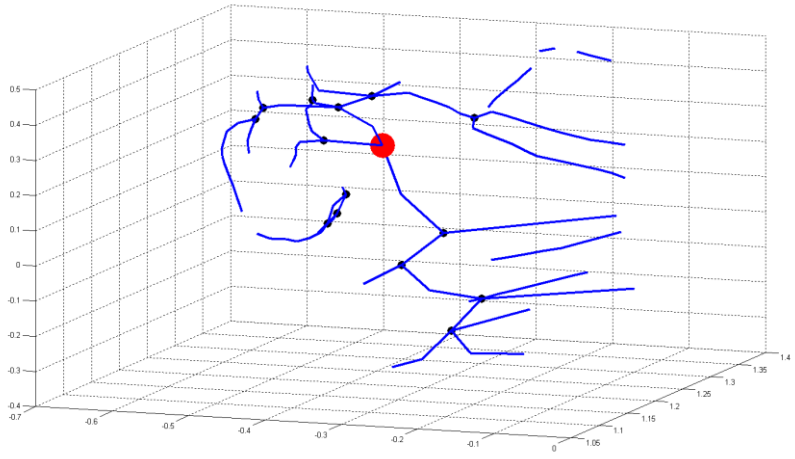


Figure 5.35 The Key Joint of the Post Cluster of Data Set Two

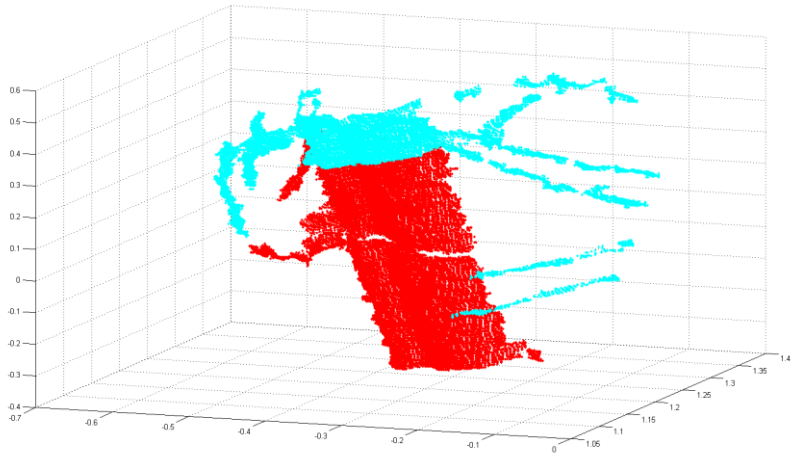


Figure 5.36 The Main Component of the Post Cluster of Data Set Two

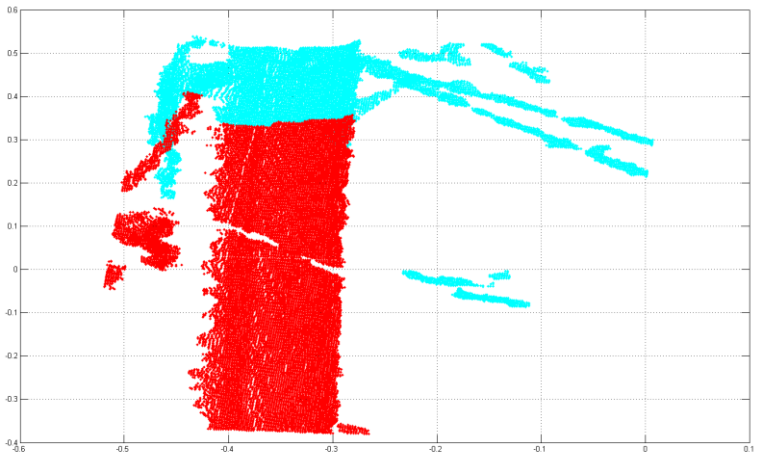


Figure 5.37 The Rotating Operation of the Post Cluster of Data Set Two

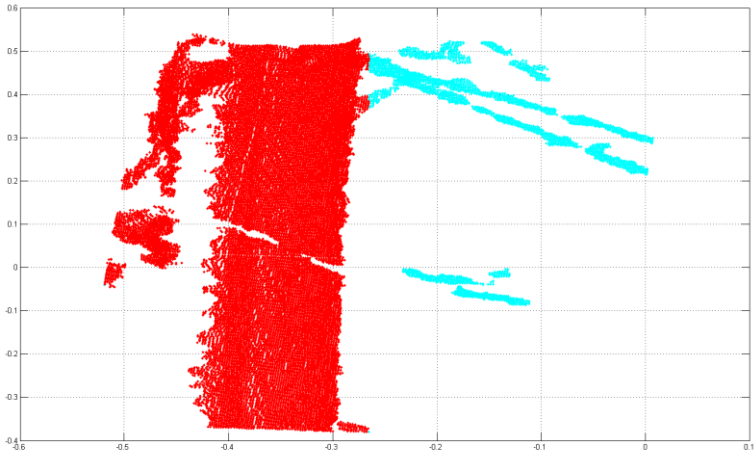


Figure 5.38 Threshold Filtering of the Post Cluster of Data Set Two

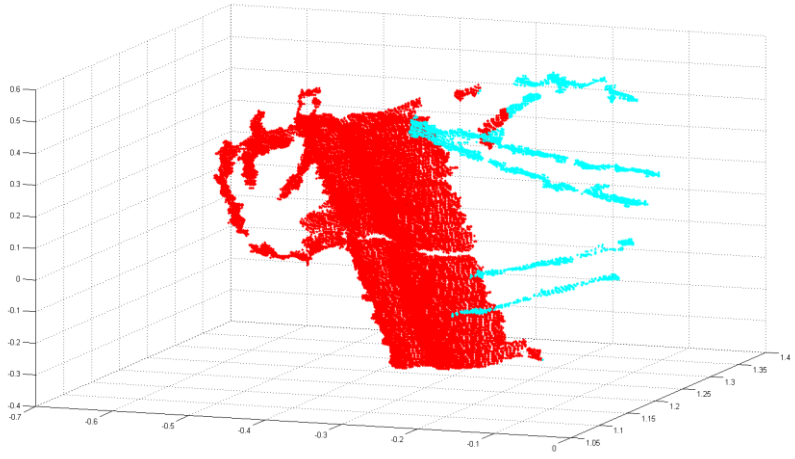


Figure 5.39 Post Identification Result of Data Set Two

After the post is extracted, the next step is to extract trunk. By using the same threshold filtering method, the trunk cluster can be directly located as shown in Figure 5.40.

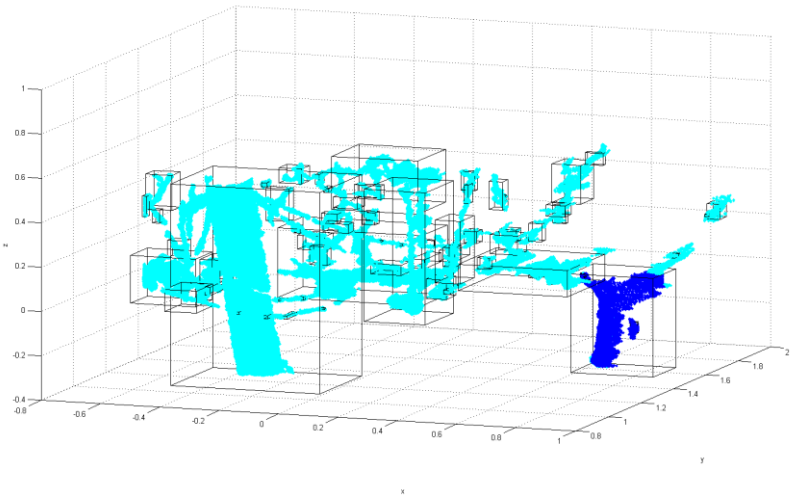


Figure 5.40 The Trunk Cluster of Data Set One (blue colour)

In order to improve the accuracy of the trunk identification, its skeleton is extracted as shown in Figure 5.41 and the key joint of the trunk is located as shown in Figure 5.42. By locating the part from the root joint to the key joint, the trunk is identified and the identification result is shown in Figure 5.43.

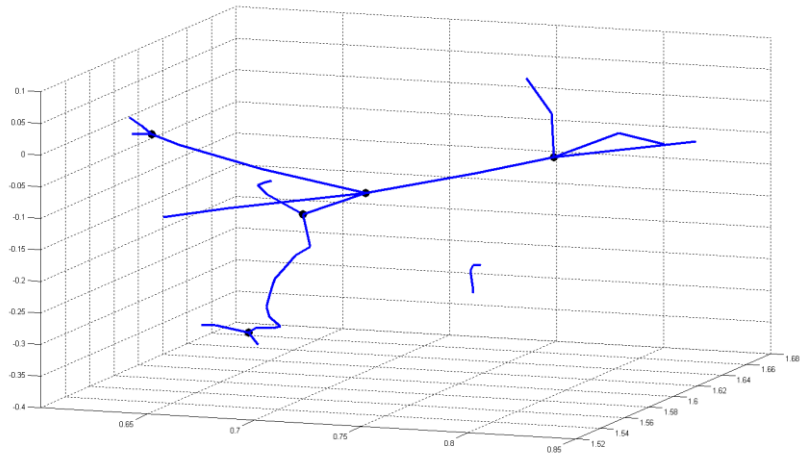


Figure 5.41 Skeleton of the Trunk of Data Set Two

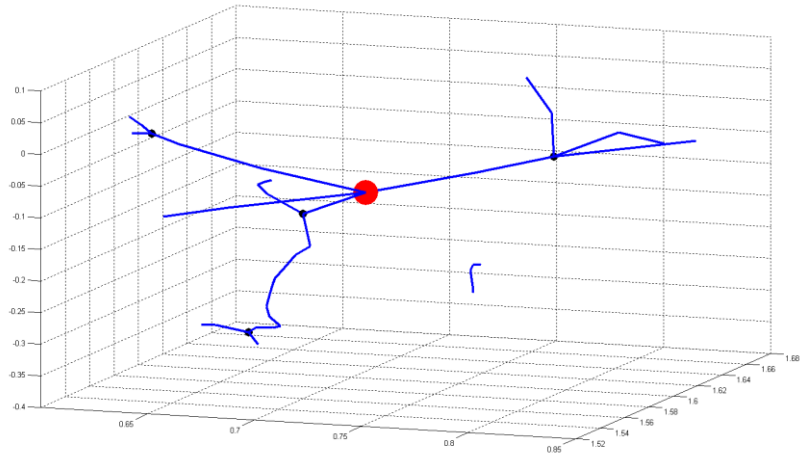


Figure 5.42 The Key Joint of the Trunk Cluster of Data Set Two

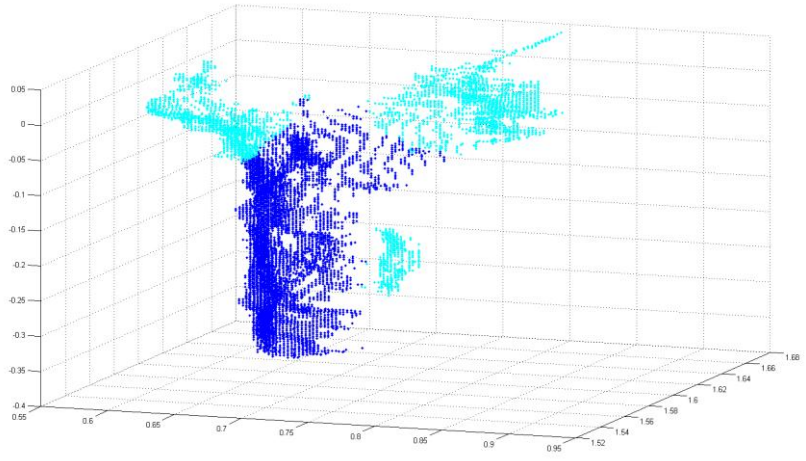


Figure 5.43 Trunk Identification of Data Set Two

After trunk extraction, the cordon identification is performed. By using the same method, the cordon clusters are extracted as shown in Figure 5.44. Unlike Data Set One, the trunk identification result of Data Set Two is refined. Therefore, it is unnecessary to keep improving the accuracy by using the threshold of cordon area.

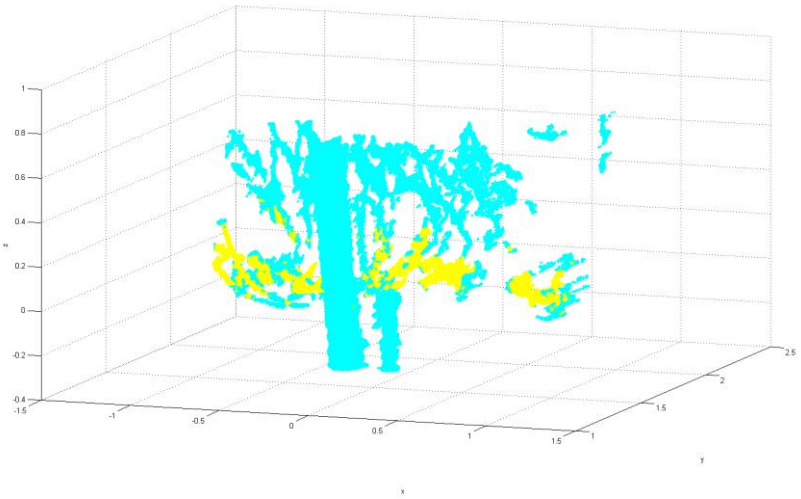


Figure 5.44 Cordon Extraction of Data Set One (yellow colour)

As there are only four objects in the pre-processed input, which are post, trunk, cordon and cane. Therefore, the left clusters are the points belonging to the cane (Figure 5.45).

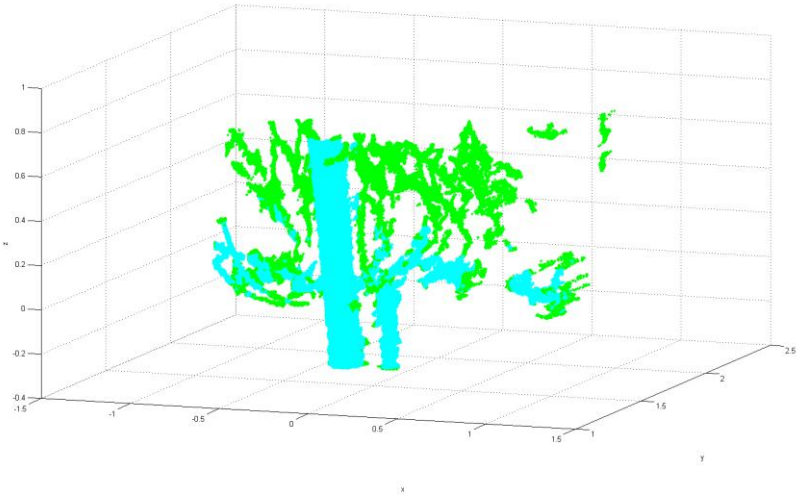
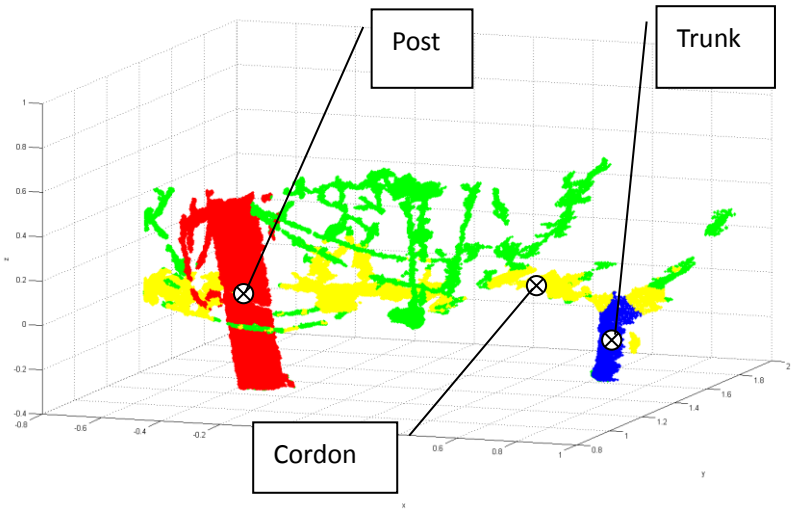


Figure 5.45 Cane Extraction of Data Set One (green colour)

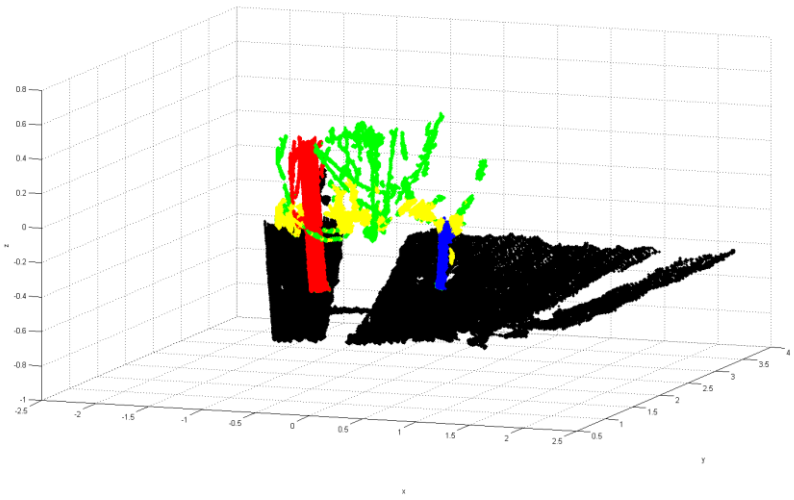
5.2.2.4 Data Set Two Final Result

The final result of the proposed method applied on data set two is illustrated in Figure 5.46. The position of post is (-0.3528, 1.2186, 0.1179). The position of trunk is (0.6629, 1.6008, -0.2788). The selected part of cordon is from 0 of x axis to 1 of x

axis and the position is (0.5020, 1.4972, 0.0167).



a. Object Identification of Data Set Two in the Pre-processed Input



b. Object Identification of Data Set Two in the Original Input

Figure 5.46 Final Result of Data Set Two

5.2.3 Data Set Three

5.2.3.1 Data Set Three Input

Data set three is the last one of the most common scenes obtained from input data sets. The difference is that only part of the post is captured in vineyards. In this case, more

than 1/2 of the post points are captured. The original input is also consisting of one post, one trunk, several cordons unconnected and scattered canes captured as shown in Figure 5.47.

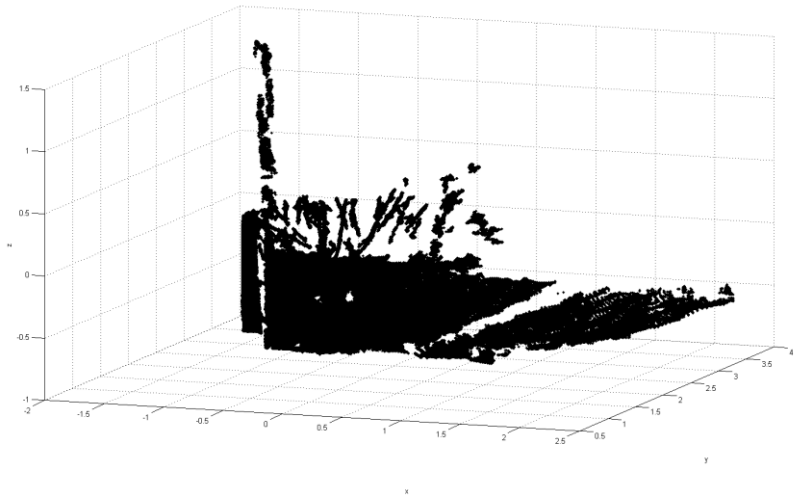
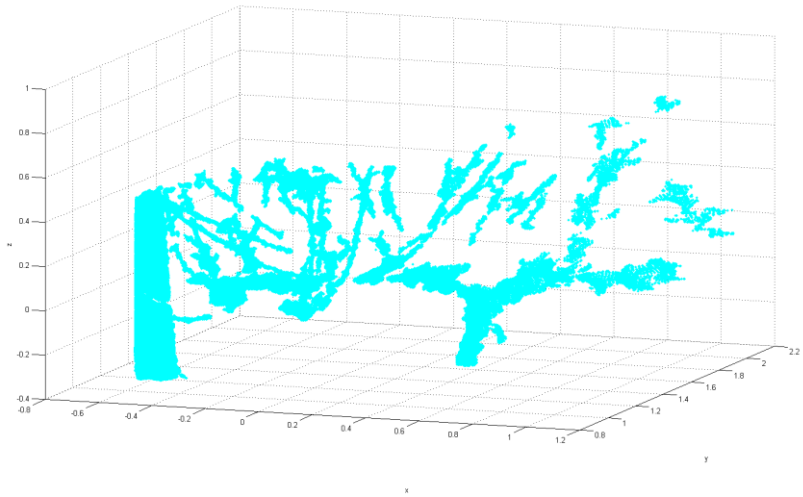


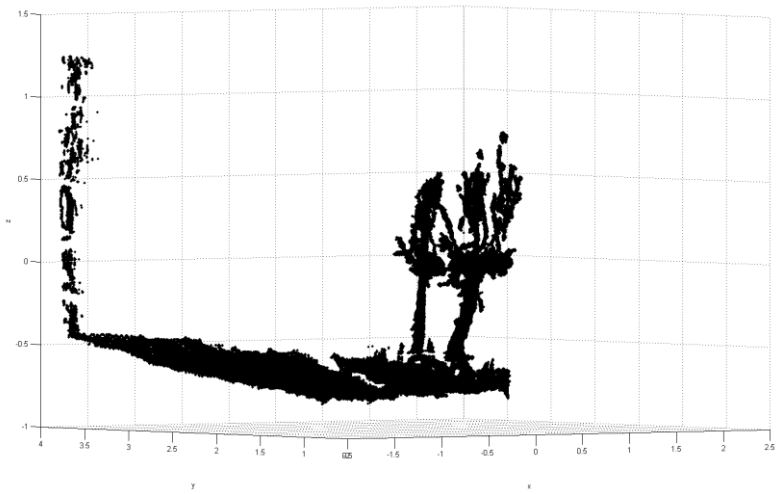
Figure 5.47 Original Input of Data Set Three

5.2.3.2 Data Set Three Segmentation

Firstly, pre-processing is performed and the result is shown in Figure 5.48-a. Not only the ground is removed but also points belonging to other rows of the grapevines as shown in Figure 5.48-b of the side view of the original input.



a. Pre-processing of Data Set Three



b. Side View of the Original Input of Data Set Three

Figure 5.48 Data Set One Pre-processing

After normal estimation, Gaussian sphere and Hough Gaussian sphere generation, the cylinder extraction is performed and the result is shown in Figure 5.49. The result of the segmentation is illustrated in Figure 5.50.

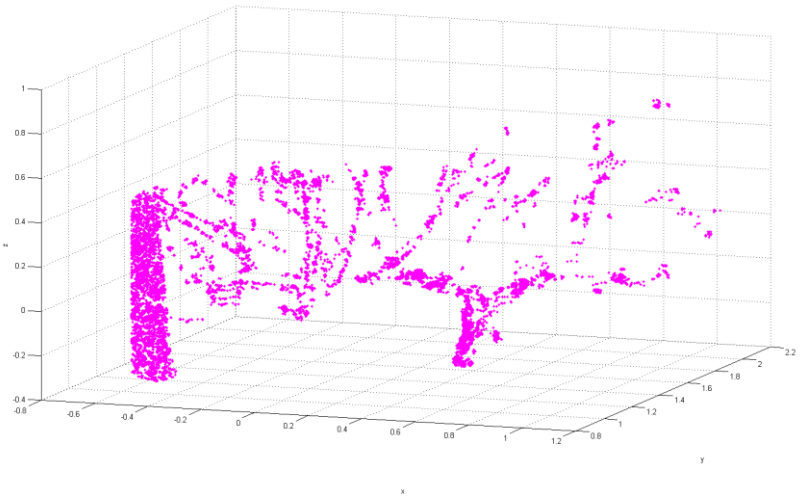


Figure 5.49 Cylinder Feature Extraction of Data Set Three

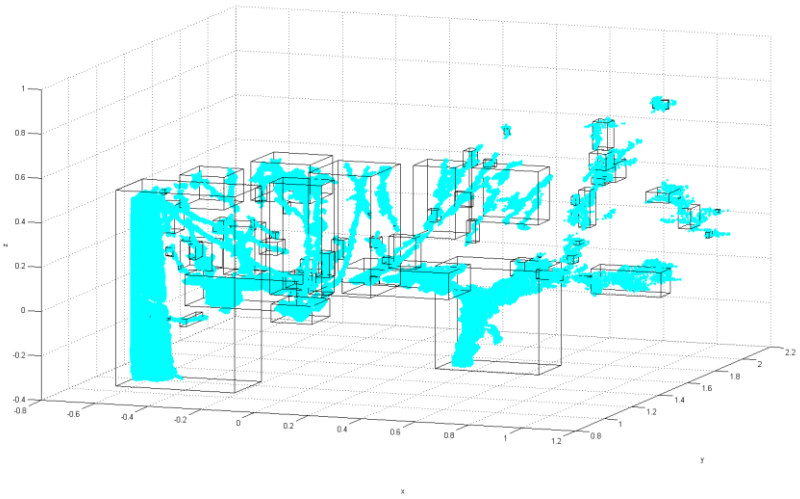


Figure 5.50 Clusters of the Pre-processed Input Points of Data Set Three

5.2.3.3 Data Set Three Objects Identification

Although only parts of the post are captured, the post cluster still contains enough number of points that fits the threshold of post cluster. The result of post cluster extraction is shown in Figure 5.51.

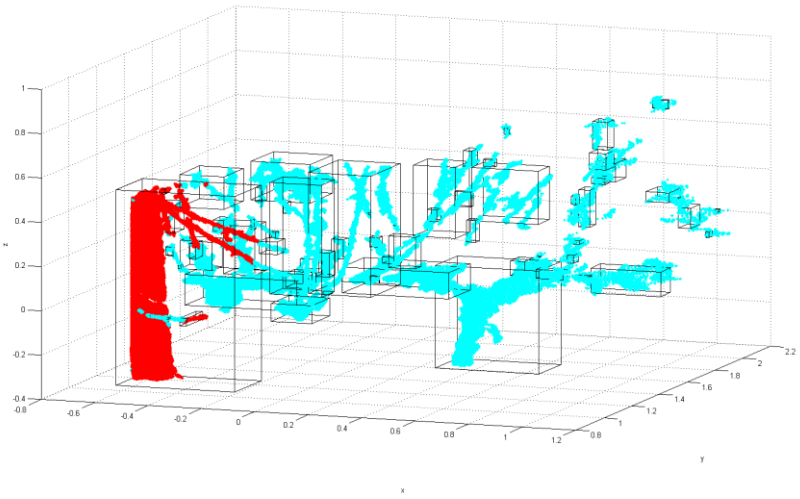
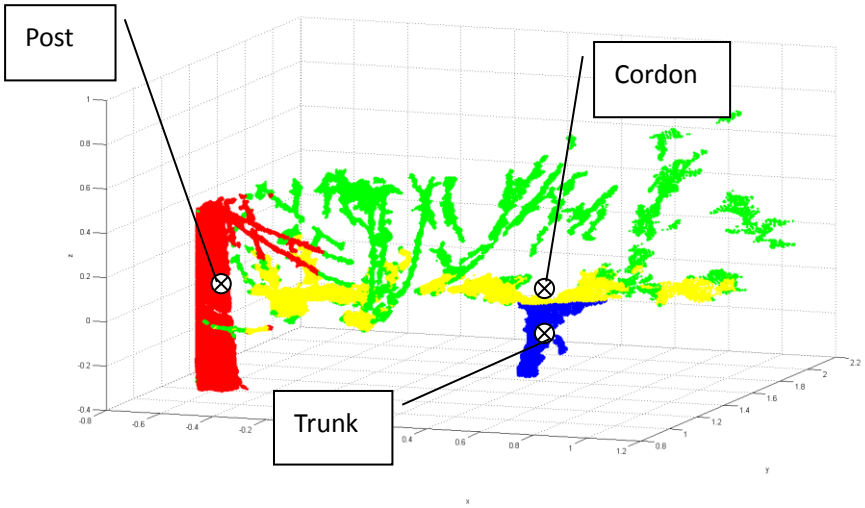


Figure 5.51 The Post Cluster of Data Set Three (red colour)

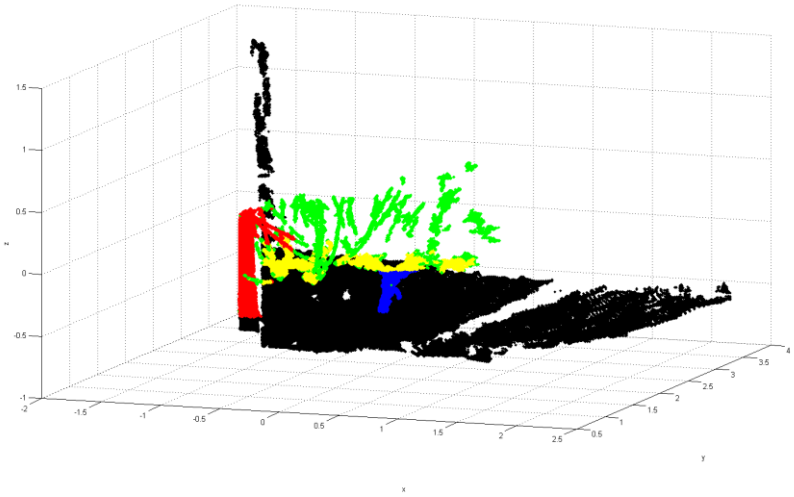
5.2.3.4 Data Set Three Final Result

The final result of the proposed method applied on data set three is illustrated in Figure 5.52. The position of post is (-0.5558, 1.1567, 0.0649). The position of trunk is

(-0.3802, 1.6400,-0.2023). The selected part of cordon is from 0.5 of x axis to 1 of x axis and the position is (0.4487, 1.6145, 0.0146).



a. Object Identification of Data Set Three in the Pre-processed Input



b. Object Identification of Data Set Three in the Original Input

Figure 5.52 Final Result of Data Set Three

5.2.4 Data Set Four

Data set four is in the same situation of data set three which is only parts of the post is captured. The only difference is that less than 1/2 of the post is captured as shown in Figure 5.53.

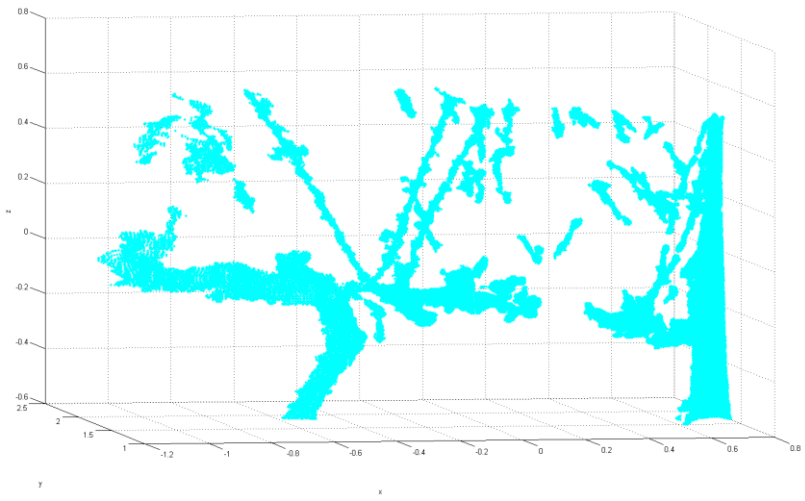


Figure 5.53 Pre-processed input of Data Set Four

Even though, the post cluster can still be extracted this demonstrates the suitability of the proposed method. The post cluster extracted is shown in Figure5.54 and the final result is shown in Figure 5.55. The position of post is (0.6472, 1.3168, -0.0442). The position of trunk is (-0.3620, 1.8246,-0.3020). The selected part of cordon is from -0.6 of x axis to -0.4 of x axis and the position is (-0.5140, 1.7745, 0.0568).

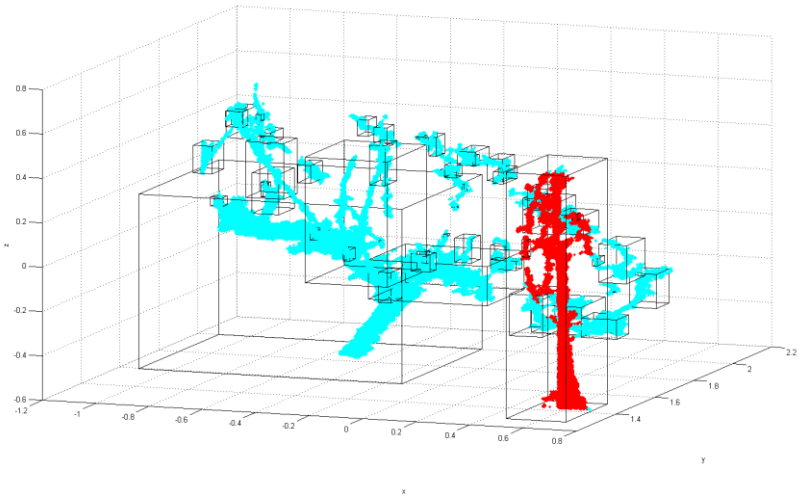
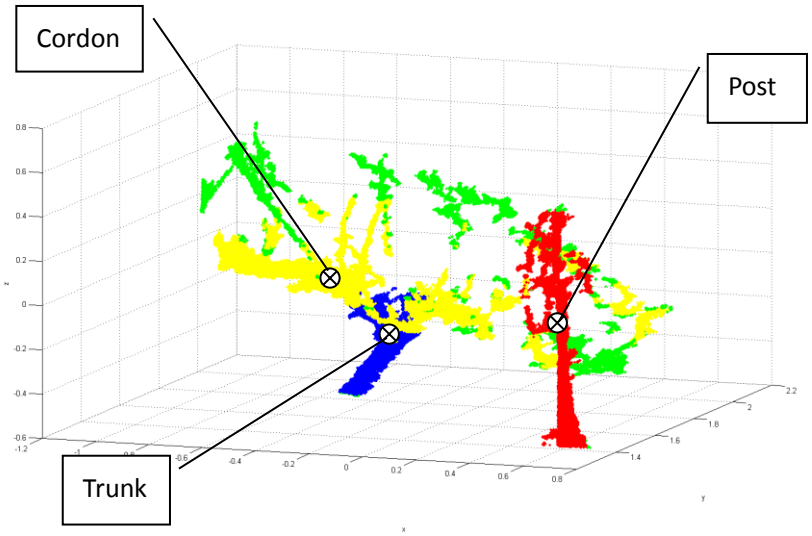
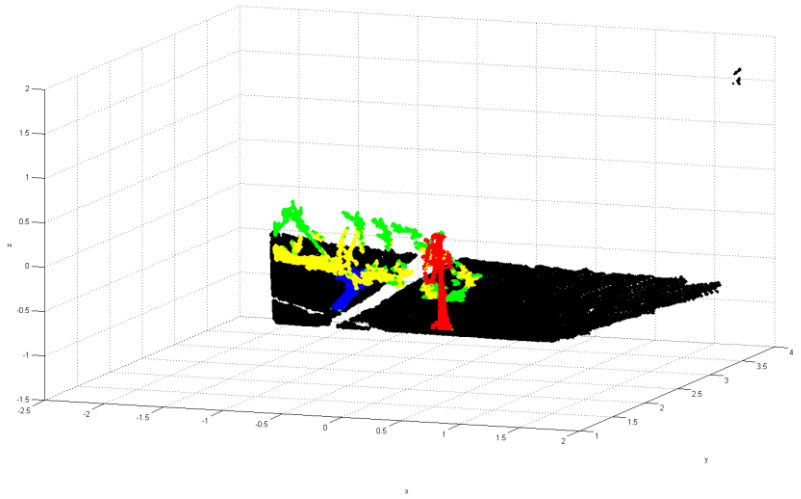


Figure 5.54 The Post Cluster of Data Set Four



a. Object Identification of Data Set Four in the Pre-processed Input



b. Object Identification of Data Set Three in the Original Input

Figure 5.55 Final Result of Data Set Four

5.2.5 Data Set Five

Data set five represents a very rare situation of the data captured in vineyards. Because of the configuration and specification of Kinect sensor and the grapevine training system, it is very unlikely to capture one post and two trunks in one scene. Under the consideration of using other range sensors and different distribution of the grapevines in different vineyards, this situation is considered and tested. As shown in Figure 5.56, the pre-processed input is consisting of the solid ground, one post and two grapevines. The proposed method is still able to extract the clusters of the two trunks captured as shown in Figure 5.57. The position of the trunk on the left is (-0.6839, 1.5403, -0.2369). The position of the trunk on the right is (0.8392, 1.7944,-0.1871).

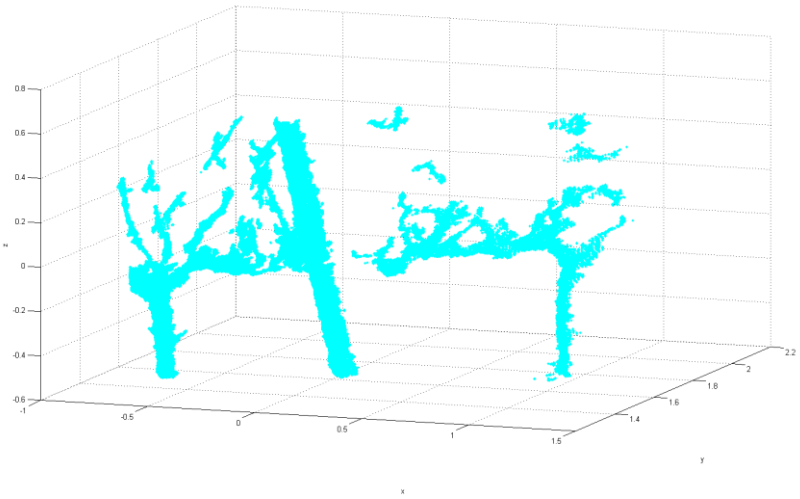


Figure 5.56 Pre-processed Input of Data Set Five

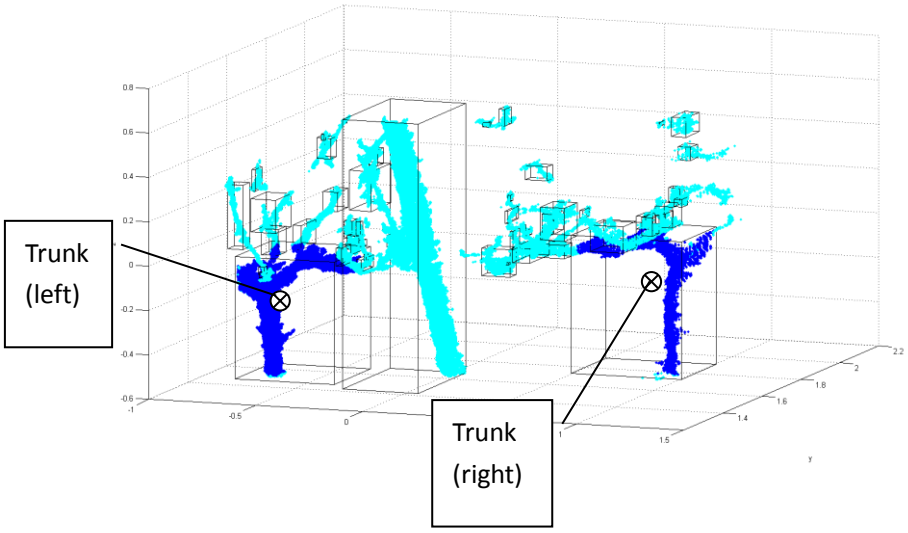


Figure 5.57 Clusters of Two Trunks of Data Set Five

5.2.6 Data Set Six

Data set six represents the other rare situation of the data captured in vineyards. The data captured contains only one grapevine which is very unlikely to happen due to the distribution of posts in vineyards. The proposed method is able to identify the objects wanted. The processed input is shown in Figure 5.58 and the final result is shown in Figure 5.59. The position of trunk is (-0.6502, 1.6699,-0.4283). The selected part of cordon is from -0.5 of x axis to 0 of x axis and the position is (-0.2662, 1.7040,

0.0919).

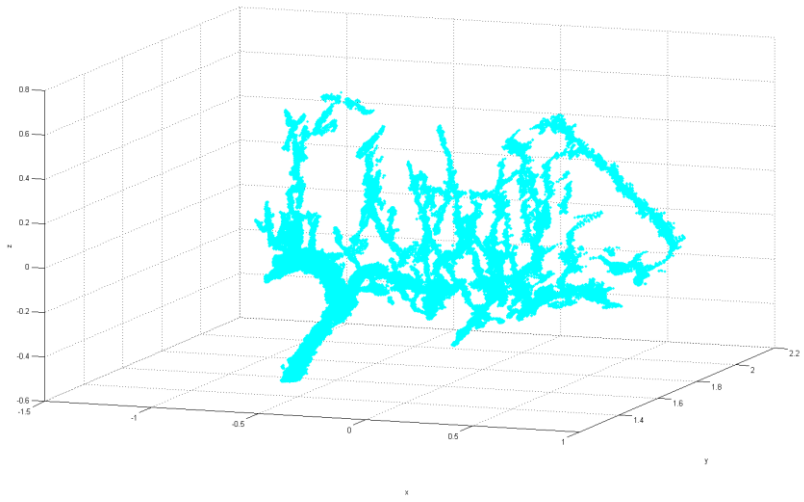


Figure 5.58 Pre-process Input of Data Set Six

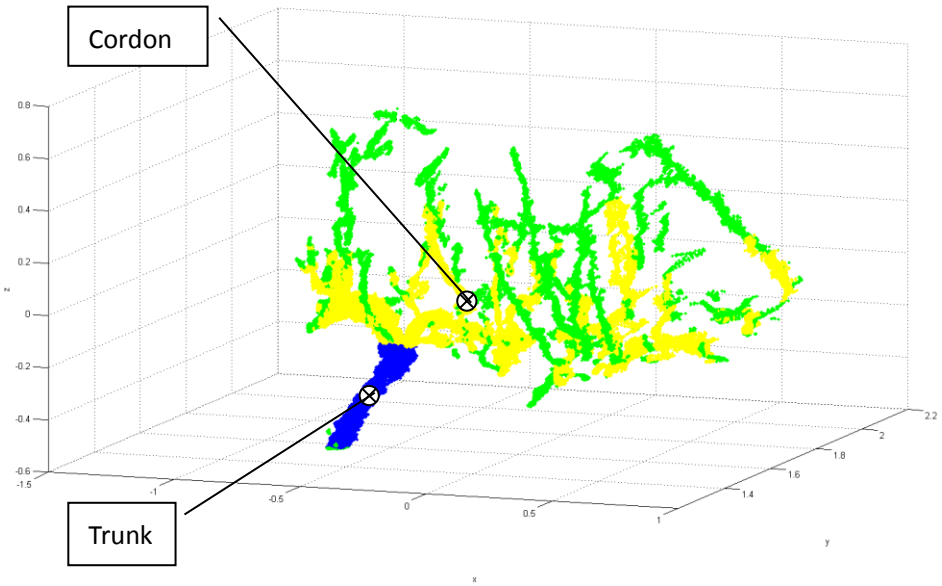


Figure 5.59 Final Result of Data Set Six

5.3 Discussion

5.3.1 Limitations

One limitation is the effectiveness of the post refinement method. In terms of data set

five which has one post and two grapevines, the post is not standing nearly vertical but with a sharp angle around 25 degrees which is a rare case as shown in Figure 5.60. From the skeleton extracted, the key joint can be obtained as shown in Figure 5.61. After the main component is extracted, the refinement result is obtained as shown in Figure 5.62. It is illustrated that the post refinement operation may produce poor results when the post is not put nearly vertical. Cases in such situation still need further investigation.

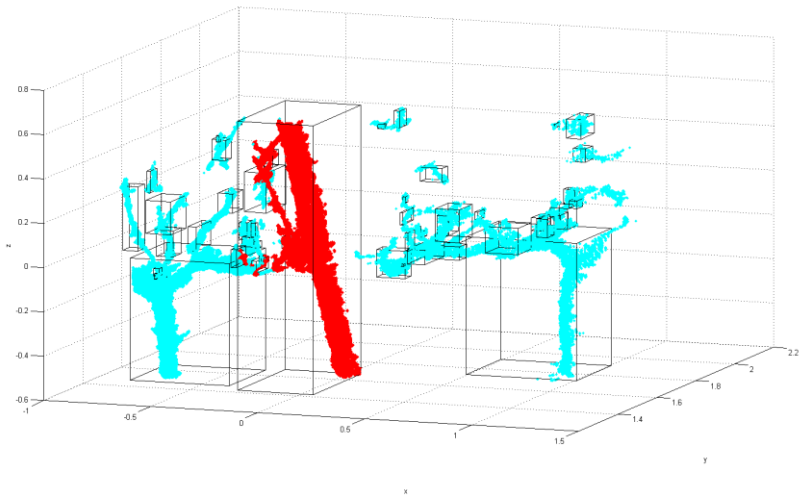
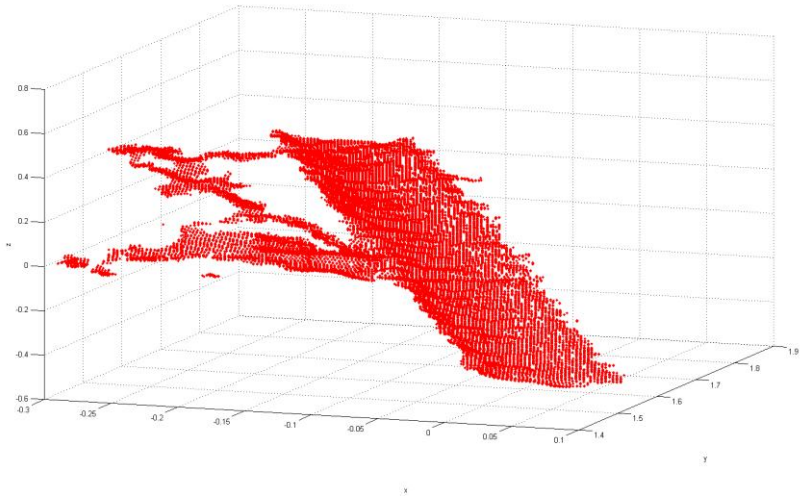
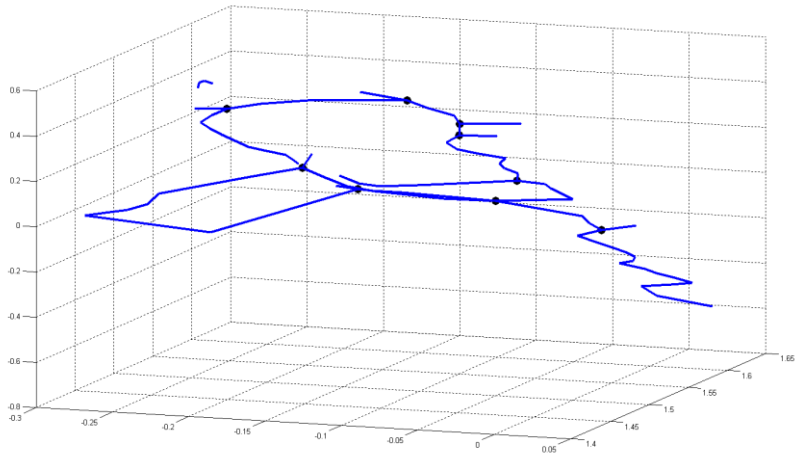


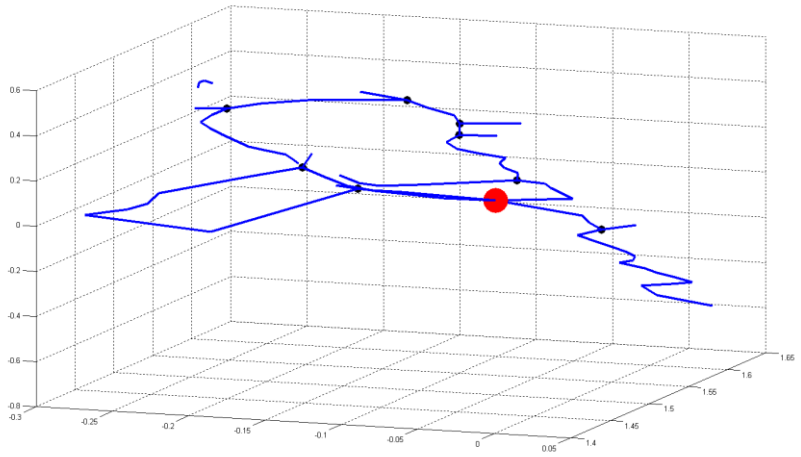
Figure 5.60 Post Cluster of Data Set Five



a. Close View of Post Cluster of Data Set Five

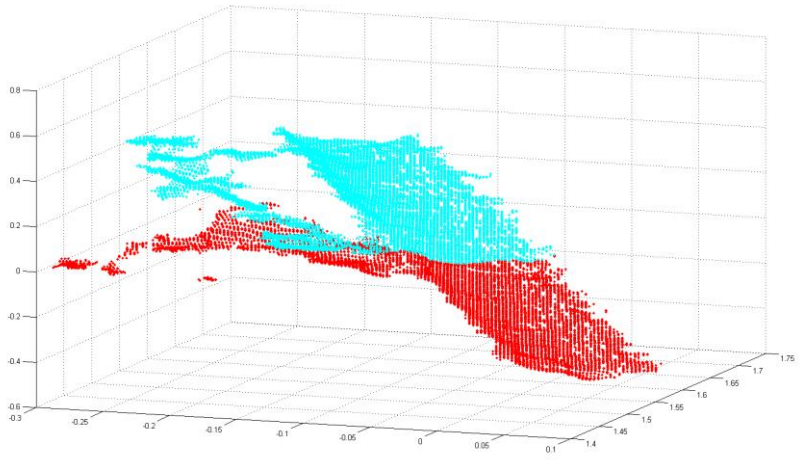


b. Post Cluster Skeleton of Data Set Five

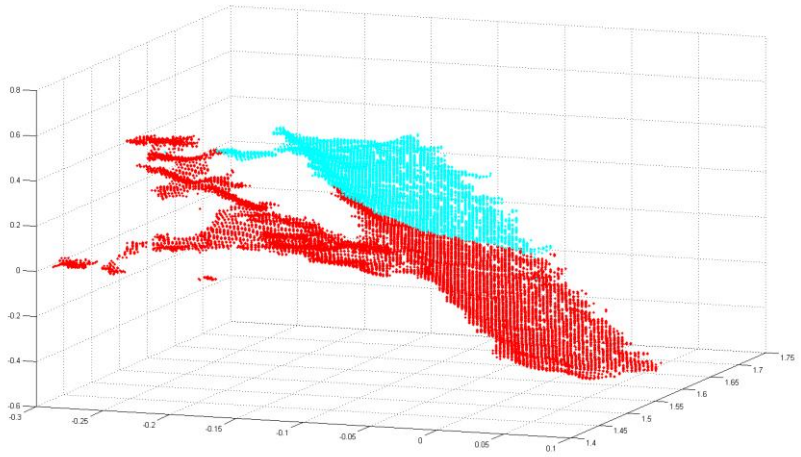


c. Key Joint of Post Cluster of Data Set Five

Figure 5.61 Skeleton Extraction of Post Cluster of Data Set Five



a. Main Component of Post Cluster of Data Set Five



b. Refinement Result of Post Identification of Data Set Five

Figure 5.62 Post Refinement of Data Set Five

The other limitation is the method is not fully autonomous. Before applying the proposed method, two input parameters are required including the cluster thresholds of post and trunk. For different vineyards and different range sensors, the threshold value, which is a number, is changed and needs to be measured via several data sets.

5.3.2 Performance

5.3.2.1 Time Consumption

The time required for each step is illustrated in Table 1. The processing speed is not fast enough for real time vineyard application.

Table 1: Time Consumption of Proposed Method

Unit - Seconds						
Data Set	Pre-processing	Normal Estimation	Hough Transform	Density Clustering	Post Skeleton	Trunk Skeleton
One	15.152353	15.152353	175.619200	18.377175	132.828236	57.868905
Two	2.388825	19.217422	161.625662	19.828163	864.642862	87.690441
Three	2.328391	16.994442	154.699218	9.757367	632.923932	73.762142
Four	2.073927	15.503144	111.393806	12.334007	177.969335	231.490486
Five	2.121518	12.078528	99.577610	12.497426	400.176914	99.986385
Six	2. 2159742	13.217676	102.440032	10.415174	N/A	167.470457

It is shown that the time spending on post and trunk skeleton extraction plays a big part of the overall time. Therefore, there is one way of improving the processing speed of the proposed method by removing the post refinement and trunk refinement steps. All the data sets are tested in such a way and data set two is taken as an example to show its result. As shown in Figure 5.63 and Figure 5.64, the result without post and trunk refinement can still work for pruning operations.

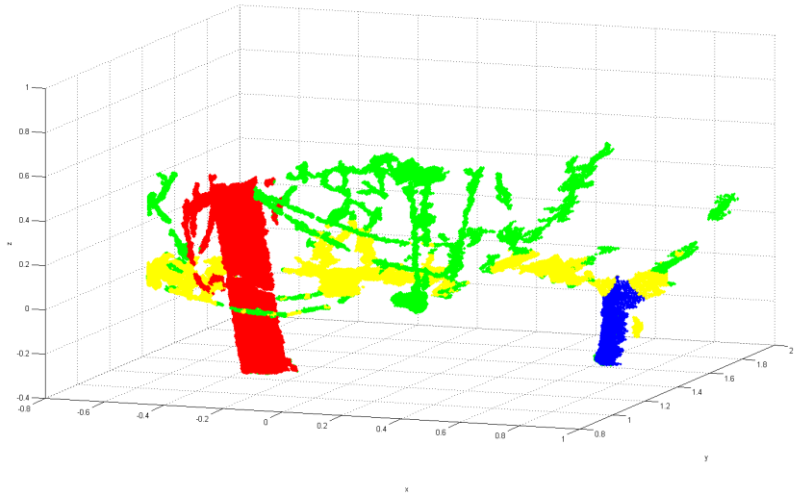


Figure 5.63 Final Result of Data Set Two with Post and Trunk Refinement

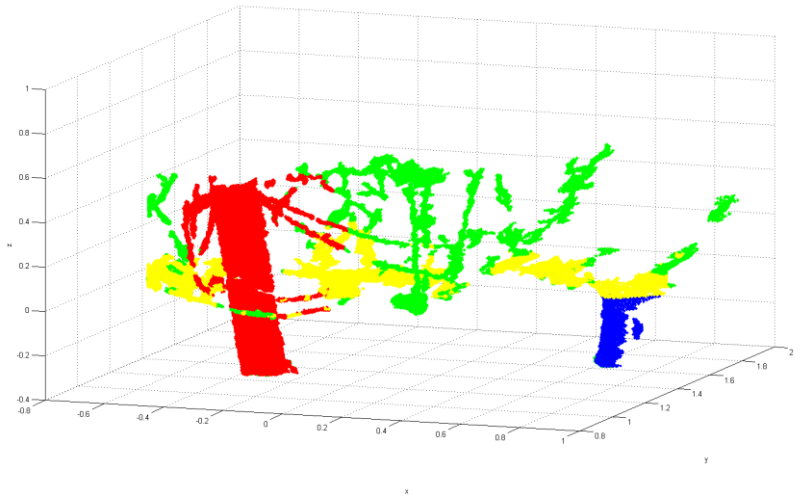


Figure 5.64 Final Result of Data Set Two without Post and Trunk Refinement

5.3.2.2 Sensitivity of the Proposed Method

There are 20 data sets tested and 6 data sets are selected to show the effectiveness and robustness of the proposed method. Among the 20 data sets, all situations in vineyards are considered and six of them are selected to analyse the performance. In terms of post and grapevine dimensions and orientations, posts of data set one and two are thicker and the whole bodies are captured; posts of data set three and four are thinner and only parts of the bodies are captured; posts of data set three and five are not set up straight; there is no post in data set six. In terms of the distance between post and grapevine, it is very short and nearly overlapped in data set one; the distance are further in data set two and three. According to the experiment results of the selected 6 data sets, the proposed method is capable of extracting post, trunk, cordon and cane under different situation which demonstrates the effectiveness of the method.

Moreover, the key concept of the method is to divide the input points into several clusters using two features synchronously, the features of cylinder and density. The only feature related to the state of the Kinect sensor is the feature of density. The density of the points extracted is not sensitive to the position change of the Kinect regarding to its distance from the grapevine, angle against the grapevine and height due to the stable performance of the Kinect sensor. Therefore, the proposed method is rubout and effective to commonly expected vineyard variations.

Chapter 6 Conclusion and Future Work

6.1 Conclusion

To achieve autonomous pruning, it is essential to identify and locate objects related to the pruning operation. As no literature has been identified in this specific area in terms of feature extraction and location derivation in vineyards, this thesis investigated and developed a method to fill this gap. A new method has been presented to automatically identify and locate four features using point cloud data from vineyards. This method applied point clouds as input to avoid the limitations of current machine vision techniques. Besides, it adopted the advantages of cylinder extraction and used density as a feature for identification. It also applied a skeleton extraction method to improve the accuracy of identification. The new idea of using cylinder extraction and density clustering synchronously for cylindrical shape objects identification was presented. As shown in the experiment results, this method is able to extract features and derive locations of the four objects from the complex scene in vineyards and is suitable for all kinds of grapevine shapes in the vineyard at the Waite Campus of the University of Adelaide.

6.2 Future Work

The Kinect sensor of Microsoft is applied to render the required point cloud data which has a fast rendering speed with enough and accurate 3 dimensional point extractions. Although the features of the Kinect are highly distinct, it is not suitable for most day time operations due to its poor performance under sunlight. As a result, a new device is desirable to be researched and developed to capture point clouds as fast and accurate as the Kinect and be functional under different illumination conditions.

Besides, the proposed method is capable of identifying post, trunk, cordon and cane and deriving their positions, but the processing speed is not fast enough for real-time

application. Algorithms and methods accordingly need to be further researched and developed in order to reduce the processing time. One of the possible ways to break through might be employing the surface feature detection on the grapevine structure so that the data points are recognized directly without many pre-processing steps.

References

1. Bartsch, T 2010, 'The cost of cane pruning in a VSP canopy', *Australian & New Zealand Grapegrower & Winemaker*, no. 558, pp. 34-36.

2. Beder, C & Förstner, W 2006, 'Direct solutions for computing cylinders from minimal sets of 3d points', *Proceedings of the 9th European conference on Computer Vision - Volume Part I*, Graz, Austria, pp. 135-146.

3. Chaperon, T & Goulette, F 2001, 'Extracting Cylinders in Full 3D Data Using a Random Sampling Method and the Gaussian Image', *Proceedings of the Vision Modeling and Visualization Conference 2001*, pp. 35-42.

4. Cao, J, Tagliasacchi, A, Olson, M, Zhang, H & Su, Z 2010, 'Point Cloud Skeletons via Laplacian Based Contraction', Proceedings of the 2010 Shape Modeling International Conference, pp. 187 - 197.

5. Daszykowski, M, Walczak, B & Massart, DL 2001, 'Looking for natural patterns in data: Part 1. Density-based approach', *Chemometrics and Intelligent Laboratory Systems*, vol. 56, no. 2, pp. 83-92.

6. Ester, M, Kriegel, H-P, Sander, J & Xu, X 1996, 'A density-based algorithm for discovering clusters in large spatial databases with noise', *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, pp. 226–231.

7. Gao, M & Lu, T-F 2006, 'Image Processing and Analysis for Autonomous Grapevine Pruning', *Proceedings of the 2006 IEEE International Conference on Mechatronics and Automation*, pp. 922-927.

8. Golub, GH & Loan, CFv 1996, *Matrix Computation*, 3rd edn, John Hopkins University Press.

9. Gümüş, B, Balaban, MÖ & Ünlüsayın, M 2011, 'Machine Vision Applications to Aquatic Foods: A Review', *Turkish Journal of Fisheries and Aquatic Sciences*, vol. 11, pp. 171-181.

10. Hoare, T 2009, 'Pruning -- cutting the cost without compromising the vineyard', *Australian Viticulture*, vol. 13, no. 3, pp. 16-18.

11. Hoppe, H, DeRose, T, Duchamp, T, McDonald, J & Stuetzle, W 1992, 'Surface reconstruction from unorganized points', *Proceedings of the 19th annual conference on Computer graphics and interactive techniques*, pp. 71-78.

12. Lukács, G, Martin, R & Marshall, D 1998, 'Faithful Least-Squares Fitting of Spheres, Cylinders, Cones and Tori for Reliable Segmentation', *Proceedings of the 5th European Conference on Computer Vision*, vol. 1, pp. 671-686.

13. Lutton, E, Maitre, H & Lopez-Krahe, J 1994, 'Contribution to the determination of vanishing points using Hough transform', *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 16, no. 4, pp. 430-438.

14. McFarlane, NJB, Tisseyre, B, Sinfort, C, Tillett, RD & Sevila, F 1997, 'Image Analysis for Pruning of Long Wood Grape Vines', *Journal of Agricultural Engineering Research*, vol. 66, no. 2, pp. 111-119.

15. Rabbani, T & Heuvel, FVD 2005, 'Efficient Hough transform for automatic detection of cylinders in point clouds', *Proceedings of the ISPRS Workshop Laser scanning*, pp. 60-65.

16. Schnabel, R, Wessel, R, Wahl, R & Klein, R 2006, 'Shape Recognition in 3D Point-Clouds', *In proceedings of The 16-th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision*.

17. Su, Y-T & Bethel, J 2010, 'Detection and Robust Estimation of Cylidner Features

in Point Clouds', *ASPRS Conference*.

18. Josep Miquel Biosca & José Luis Lerma 2008, ' Unsupervised robust planar segmentation of terrestrial laser scanner point clouds based on fuzzy clustering methods ', '*ISPRS Journal of Photogrammetry and Remote Sensing*', vol.63, no.1, pp.84–98.

19. Sergey Tyrin & Itamar Barkai 2009, ' Automatic Grape Clusters Detection in Vineyard Images'，*Final Project of Computer Science, Ben Gurion University of the Negev.*

20. Tim Braun, Heribert Koch, Oliver Strub, Gregor Zolynski & Karsten Berns 2010, ' Improving Pesticide Spray Application in Vineyards by Automated Analysis of the Foliage Distribution Pattern in the Leaf Wall ', *CVT 2010 - March 16-18, Kaiserslautern, Germany.*

21. Christopher Weber, Stefanie Hahmann & Hans Hagen 2010, 'Sharp Feature Detection in Point Clouds', *In Proceedings of the 2010 Shape Modeling International Conference (SMI '10), IEEE Computer Society, Washington, DC, USA*, pp.175-186.

22. Park, MK, Lee, SJ & Lee, KH 2012, 'Multi-scale Tensor Voting for Feature Extraction from Unstructured Point Clouds', *Graphical Models*, vol. 74, no. 4, pp. 197-208.

23. Holies, RC & Fischler, MA 1981, 'A RANSAC-based Approach to Model Fitting and its Application to Finding Cylinders in Range Data', *Proceedings of the 7th international joint conference on Artificial intelligence, Vancouver, BC, Canada*, vol. 2, pp. 637-643.

24. Lozano-Perez, T, Grimson, W & White, S 1987, 'Finding cylinders in range data', *Robotics and Automation. Proceedings. 1987 IEEE International Conference on,*

vol. 4, pp. 202-207.

25. Gorte, B 2006, 'Skeletonization of Laser-Scanned Trees in the 3D Raster Domain Innovations in 3D Geo Information Systems', *in A Abdul-Rahman, S Zlatanova & V Coors (eds), Springer Berlin Heidelberg,* pp. 371-380.

26. Gorte, B & Pfeifer, N 2004, 'Structuring Laser-scanned Trees using 3d Mathematical Morphology', *International Archives of Photogrammetry and Remote Sensing*, vol. 35, pp. 929–933.

27. Pfeifer, N, Gorte, B & Winterhalder, D 2004, 'Automatic Reconstruction of Single Trees from Terrestrial Laser Scanner Data', *Proceedings of 20th ISPRS Congress*, pp. 114-119.

28. Bucksch, A & Lindenbergh, R 2008, 'CAMPINO — A skeletonization method for point cloud processing', *SPRS Journal of Photogrammetry and Remote Sensing*, vol. 63, no. 1, pp. 115-127.

29. Bucksch, A, Lindenbergh, RC & Menenti, M 2009, 'SkelTre - Fast Skeletonisation for Imperfect Point Cloud Data of Botanic Trees', *Eurographics Workshop on 3D Object Retrieval (3DOR'09)*.

30. Livny, Y, Yan, F, Olson, M, Chen, B, Zhang, H & El-Sana, J 2010, 'Automatic reconstruction of tree skeletal structures from point clouds', *ACM Trans Graph*, vol. 29, no. 6, pp. 1-8.

31. Linda Shapiro &George Stockman 2001, 'Computer Vision', *Prentice-Hall Inc.*

32. Augusto Sarti & Stefano Tubaro 2002, 'Detection and Characterisation of Planar Fractures using a 3D Hough Transform', *Signal Processing*, vol.82, no.9, pp. 1269–1282.