DETECTION OF CHERRY TREE BRANCHES AND LOCALIZATION OF SHAKING POSITIONS FOR AUTOMATED SWEET CHERRY HARVESTING

By

SURAJ AMATYA

A dissertation submitted in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

WASHINGTON STATE UNIVERSITY
Department of Biological Systems Engineering

DECEMBER 2015

© Copyright by SURAJ AMATYA, 2015
All Rights Reserved
To the Faculty of Washington State University:

The members of the Committee appointed to examine the thesis of SURAJ AMATYA find it satisfactory and recommend that it be accepted.

Manoj Karkee, Ph.D., Chair

Qin Zhang, Ph.D.

Matthew D. Whiting, Ph.D.
ACKNOWLEDGEMENT

I would like to express my sincere gratitude to my advisor Dr. Manoj Karkee for the continuous support for my PhD research and for his patience, motivation and encouragement throughout my course of study. His guidance helped me to overcome all the hurdles that I faced during my research. I could not have completed my dissertation without his mentoring.

Besides my main advisor, I would like to thank my committee members Dr. Qin Zhang and Dr. Matthew Whiting for their insightful comments and encouragement.

I am also grateful to my colleagues at CPAAS especially Aleana Gongal, Peter Larbi, Jianfeng Zhou, Long He, Patrick Scharf, Yunxiang Ye, Mark DeKleine, Abhisesh Silwal and Aadit Shrestha for their help and support in conducting my research and field work.
Sweet cherry industry has been challenged by increasingly uncertain availability and rising cost of farm labor. As harvesting is the most labor-intensive operation, there is a need for developing automated harvesting solutions to address labor related challenges. Over the past several decades, researchers have studied mechanical shakers for efficient sweet cherry harvesting. Automating a shaking harvester using a machine vision system for detecting and locating tree branches has the potential to further reduce the labor demand for cherry harvesting.

This research focused on detecting cherry tree branches in full foliage canopies. The branch detection method used visible branch sections in canopy images to reconstruct whole branches. Morphological features of the visible branch sections including orientation, length and thickness were used to group sections of the same branch together, which were then connected using a model equation. This method achieved a branch detection accuracy of 89% in cherry trees trained in vertical trellis system.
Using this method, in Y-trellis system with denser foliage, only 55% branches were detected. Therefore, cherry clusters were detected and integrated for locating highly occluded branches. This method identified series of cherry clusters growing along specific branches and a model equation was fitted to reconstruct complete branches. This method led to a branch detection accuracy of 94% in Y-trellis system.

After branch detection, a method was developed to determine shaking positions on branches for effective cherry harvesting. Shaking locations were first identified in the color images. For each shaking position identified in the color image, 3D location information was estimated using 3D camera images. The distance to shaking positions from the camera was estimated with a root mean square value of 0.064 m with this method. Cherry tree branches were then shaken at the estimated locations using a handheld shaker. Maximum fruit removal efficiency was found to be 93% and 87% in Y- and vertical trellis system respectively. The results of this research have shown a great potential for machine vision-based automated solutions for cherry harvesting.
TABLE OF CONTENTS

ACKNOWLEDGEMENT .................................................................................................................. iii

ABSTRACT ....................................................................................................................................... iv

TABLE OF CONTENTS .................................................................................................................. vi

LIST OF FIGURES ........................................................................................................................ x

LIST OF TABLES ........................................................................................................................ xv

CHAPTER ONE: INTRODUCTION ............................................................................................... 1

1.1 Background ............................................................................................................................. 1

1.2 Literature Review .................................................................................................................. 3

1.2.1 Tree Fruit Harvesting ...................................................................................................... 3

1.2.2 Current Status of Sweet Cherry Harvesting ................................................................. 6

1.3 Goals and Objectives ............................................................................................................ 8

1.4 Relevant Previous Research ............................................................................................... 9

1.5 Dissertation Organization ................................................................................................. 10

References ...................................................................................................................................... 11

CHAPTER TWO: FUNDAMENTAL MACHINE VISION CONCEPTS ........................................... 17

2.1 Digital Image Acquisition ................................................................................................. 18
CHAPTER FOUR: INTEGRATION OF VISIBLE BRANCH SECTIONS AND CHERRY CLUSTERS FOR DETECTING CHERRY TREE BRANCHES IN DENSE FOLIAGE CANOPIES

4.1 Introduction

4.2 Materials and Methods

4.2.1 Image Acquisition

4.2.2 Image Classification

4.2.3 Branch Detection using Branch Pixels

4.2.4 Combining Branch and Cherry Regions for Improved Branch Trajectory

4.2.5 Cherry Pixel-Based Branch Detection

4.3 Results and Discussions

4.3.1 Pixel Classification

4.3.2 Tree Branch Detection

4.4 Conclusion

References

CHAPTER FIVE: A METHOD FOR LOCALIZING SHAKING POSITIONS IN CHERRY TREE BRANCHES FOR AUTOMATED SWEET CHERRY HARVESTING
5.1 Introduction

5.2 Materials and Methods

5.2.1 Test Orchard

5.2.2 Image Acquisition

5.2.3 Co-registration of Depth and RGB Images

5.2.4 Branch Detection and Reconstruction

5.2.5 Determining Shaking Locations in Tree Branches

5.3 Results and Discussions

5.3.1 Mapping 3D depth information onto RGB image

5.3.2 Fruit Removal Efficiency

5.4 Conclusion

References

CHAPTER SIX: CONCLUSIONS AND RECOMMENDATIONS

6.1 General Conclusions

6.2 Recommendations for Future Work
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1.1</td>
<td>Cherry harvesting practices in the past and the present</td>
</tr>
<tr>
<td>Figure 2.1</td>
<td>An example application of machine vision system in apple orchard</td>
</tr>
<tr>
<td>Figure 2.2</td>
<td>Digital image acquisition process; energy emitted by illumination source is reflected by scene elements, which is captured by the imaging sensor and recorded as a digitized image (Gonzalez and Woods, 2002)</td>
</tr>
<tr>
<td>Figure 2.3</td>
<td>RGB color model; additive colors: red and green produce yellow, green and blue produce cyan, red and blue produce magenta</td>
</tr>
<tr>
<td>Figure 2.4</td>
<td>a) Hue, Saturation, and Lightness (HSL) color model and b) Hue, Saturation, and Value (HSV) color model (Source: Wikipedia)</td>
</tr>
<tr>
<td>Figure 2.5</td>
<td>Gray-level histograms depicting opportunities for image segmentation by single threshold (left) and multiple thresholds (right) (Gonzalez and Woods, 2002)</td>
</tr>
<tr>
<td>Figure 2.6</td>
<td>Probability distribution of one dimensional feature vector for class $w_1$ and $w_2$</td>
</tr>
<tr>
<td>Figure 2.7</td>
<td>Working principle of time-of-flight-of-light (TOF) based camera (Shim and Lee, 2012)</td>
</tr>
<tr>
<td>Figure 2.8</td>
<td>Photonic-Mixer-Device (PMD) camera based on time-of-flight-of-light (TOF)</td>
</tr>
<tr>
<td>Figure 2.9</td>
<td>RGB camera (Bumblebee XB3, top) and 3D camera (PMD CamCube 3.0, bottom) pair used for mapping depth and color information</td>
</tr>
</tbody>
</table>
Figure 3.1: Simplified representation of a cherry tree canopy trained in upright fruiting offshoots (UFO) system ................................................................. 38

Figure 3.2: Field image acquisition setup with artificial lighting. Camera mounted on a pan-n-tilt system took three images capturing region between wire 1 and wire 4 .................. 40

Figure 3.3: a) An image of cherry tree; and b) image after image enhancement by increasing contrast ratio ........................................................................................................... 41

Figure 3.4: a) Original image with specular reflections on cherry centres and along medial axis of branches; and b) image after removal of specular reflections by inward interpolation method .......................................................... 42

Figure 3.5: Equivalent ellipse, major axis length and minor axis length of a segmented region (shaded region) ........................................................................................................... 45

Figure 3.6: Flowchart of the branch detection method; The input to the process is a segmented cherry tree image and the output is the number of branches in an image and equations representing those branches .................................................. 46

Figure 3.7: Skeletonization and fragmentation of cherry branch segments; Medial axis skeleton was obtained and branching points were identified in the skeleton; Branches were then divided at branching points to form smaller, individual sections ........................................ 48

Figure 3.8: Orientations of branch segments represented by arrows; Two regions were assumed to be parts of the same branch when their relative orientations were within 20°........................................ 50

Figure 3.9: Decision boundaries for different classes of pixels in R-G, R-B, and B-G planes of a RGB colour space ............................................................... 52
Figure 3.10: Sample result of Bayesian classification of an image to branch, cherry, leaf and background classes ................................................................. 56

Figure 3.11: Individual branch detection: original image (left); detected branches (right) .......... 59

Figure 4.1: Sensing system used in the study to acquire images of cherry trees trained in Y-trellis architecture. Image acquisition was carried out in night time using LED lights (Trilliant ® 36, Grote). ........................................................................................................ 73

Figure 4.2: a) Segmented branch regions (gray); b) Detection of branch; striped regions represented detected branch segments; gray regions represent branch regions not detected as a part of any branch; c) improved branch detection by combining neighborhood branch regions to form a single branch ........................................................................ 76

Figure 4.3: a) Segmented branch region (gray); b) Detection of branch; horizontally striped regions represent detected branch segments; gray regions are segmented cherry regions; c) improved branch detection by combining cherry regions in the neighborhood of previously detected region ................................................................. 77

Figure 4.4: Flowchart showing major steps in branch detection using branch and cherry pixel regions ............................................................................................................. 79

Figure 4.5: A neighborhood for identifying cluster of cherries around a given cluster for predicting a branch passing through them ......................................................... 81

Figure 4.6: Progression of the region growing method with a neighborhood search around a cherry region; a) first step from bottom-most region with no offset in x-axis; b) second step with an x-axis offset of (x2-x1)/2; and c) third step with x-axis offset of (x3-x2)/2 .......... 81
Figure 4.7: An example result of the branch detection method; a) original image of a cherry tree; and b) identified branches and their respective equations. Square markers represent branches identified using branch pixels, and circular markers represent branches identified using cherry pixels.

Figure 5.1: a) Upright Fruiting Offshoots (UFO) vertical planar architecture; and b) UFO Y-trellis architecture.

Figure 5.2: Geometric representation of co-registration of 3D and 2D images (Gongal et al., 2015).

Figure 5.3: Flowchart of the branch detection algorithm.

Figure 5.4: A sample cherry tree image divided into three zones to locate necessary shaking positions for each branch.

Figure 5.5: Flow chart of the shaking point localization process in tree canopies.

Figure 5.6: a) An example zone of a cherry tree branch showing clusters of cherries and a visible branch segment; b) Segmented branch region (white region), branch trajectory defined by the corresponding branch equation (black-dotted line). The median position of overlapping co-ordinates between branch region and branch trajectory was selected as the shaking position.

Figure 5.7: a) An example zone of a cherry tree branch showing clusters of cherries and a visible branch segment; b) A scenario when branch equation does not pass through branch region; red dotted line represent the path defined by branch equation, segmented branch and cherry regions are shown as gray and blue striped regions.
Figure 5.8: A part of a cherry tree canopy where the branch section has been completely occluded by fruit and leaves; dashed line represents the trajectory defined by the branch equation and the circle below the largest cherry cluster represents the shaking position. 107

Figure 5.9: a) Original cherry branch image with overlapping branch trajectory (dotted white); b) segmented cherry region (white) and branch trajectory (black dotted); c) cherry distance profile showing distance to nearest cherry region for each coordinate in branch path 110

Figure 5.10: Secondary shaking position determination for second iteration of harvesting; a) Potential shaking positions selected as P1, P2, P3, & P4, which were 10 pixels away from nearest cherry cluster; b) Impact region of P1; c) Impact region of P2; and d) Impact region of P3 111

Figure 5.11: Error in estimating distance to branch sections through mapping of 3D information onto RGB images. Solid line represents mean whereas shaded band represent standard deviation region 113

Figure 5.12: a) Fruit removal efficiency (percentage) for shaking at each consecutive position per branch on Y-trellis canopy system; and b) Cumulative fruit removal efficiency after each consecutive shaking on Y-trellis canopy system 115

Figure 5.13: a) Fruit removal efficiency (percentage) for shaking at each consecutive position per branch on vertical trellis canopy system; and b) Cumulative fruit removal efficiency after each consecutive shaking on vertical canopy system 116

Figure 5.14: Number of shaking positions required for maximum fruit removal in Y-trellis system (a) and in vertical trellis system (b). 117
LIST OF TABLES

Table 3.1: Classification accuracy for two-, three-, and four-class classification with training dataset. ................................................................................................................................... 54

Table 3.2: Classification accuracy for two-, three-, and four-class classification with testing dataset. ................................................................................................................................... 55

Table 3.3: Confusion matrix of testing set for four class classification........................................ 55

Table 3.4: Accuracy in detecting individual cherry branches....................................................... 58

Table 4.1: Mean feature vector for branch, cherry, leaf and background classes......................... 83

Table 4.2: Confusion matrix depicting the performance of pixel classification method.......... 83

Table 4.3: Results of branch detection method based on branch pixels and cherry pixels......... 85

Table 5.1: Error in estimating distance to shaking positions from the 3D camera. Distance was estimated by mapping 3D depth information onto color images................................. 113
CHAPTER ONE

INTRODUCTION

1.1 Background

For more than a century, the advancement in agricultural technologies has been shaping the ways people grow food. From the complete reliance on animal power and human muscle for farming, the farmers have now embraced machines for most of the agricultural operations. Horses and mules used for agriculture in earlier days have now been replaced by over 5 million tractors in US (Dimitri et al., 2005). For example, development of combine harvesters has significantly reduced the labor required for harvesting grain crops. A crew of four men with a 15-foot combine, two operating the combine and two hauling grain, harvested more than 35 acres a day, which would have required 15 to 16 men with headers and a stationary thresher and 18 to 20 men with binders (The mechanization of American agriculture, 1927).

Today’s agriculture is generally characterized by a small number of large, specialized farms, which are highly productive and mechanized. Past a few decades have shown a huge leap towards automated and robotic technologies for agriculture including GPS/vision guided machinery systems such as tractors, planters, combine harvesters, and robotic orchard platforms. With the advancement of sensing, communication, and control technologies coupled with Global Navigation Satellite System (GNSS) and Geographic Information System (GIS), the agricultural machines are being transformed from simple, mechanical machines of yesterday to the intelligent, autonomous vehicles of the future (Zhang and Pierce, 2013).
In spite of all these technological advancements, the way tree fruit crops are harvested for fresh market consumption has not moved much over the last century (Figure 1.1). Developing mechanical harvesters for tree fruit presents great complexity as fruit are located high above the ground at random spatial locations (Li et al., 2011). There are several complexities in automating tree fruit harvesting including adaptation to horticultural characteristics, variation in terrain, climate, and fruit maturity, as well as complication in substituting human judgement and dexterity (Sarig, 2005).

*Figure 1.1: Cherry harvesting practices in the past and the present*

(Source: http://maraschinocherries.org/the-growers/history-of-cherry-growing-the-us/)

*Cherry harvesting in 1950s*

(Source: http://ijpr.org/post/some-washington-cherry-orchards-short-labor-year [left];

*Cherry harvesting in 2015*
Hand picking of fruit involves retrieving fruit from random spatial locations on the tree canopies (Li et al., 2011) into a basket and transferring it to collection bins. In most cases, the pickers need to climb up a ladder to retrieve fruit from high canopy regions and descend with heavy loads of harvested fruit. Such harvesting activities expose the fruit pickers to high risks of back strain, and musculoskeletal problems (Fathallah, 2010). In addition, fall and other accidents related to ladder use are prevalent (Hoffman et al., 2006). It has been reported that $21 million was paid in compensation for ladder-related injuries by the Washington State tree fruit industry between 1996 and 2001 (Hofmann et al., 2006). Due to the increasing labor cost and uncertainty of labor availability, it is expected that the labor-related issues are going to be more critical in the near future. Automated harvesting has the potential to address this issue but requires advancement in several areas. Some of the major problems to be solved include recognizing and localizing canopy objects including fruit and branches, effectively detaching fruit from the tree, and keeping the harvest-induced damage to tree and fruit to an acceptable level (Sarig, 1992).

1.2 Literature Review

1.2.1 Tree Fruit Harvesting

As discussed in Sec 1.1, commercially adopted tree fruit harvesting methods have remained largely the same over the past 100 years. Today, almost 100% of fresh market tree fruit crops around the world is harvested in a conventional way by human pickers. Although the need for mechanized or robotic harvesting has been realized and decades of research have been conducted in this area, the commercial success of such harvesters has been limited. In industrial automation and robotic applications, the work environment is designed specifically for the robot’s optimal performance eliminating as many variables as possible. However, in agricultural settings,
environmental, weather and crop-related variables pose significant hurdle to successful automation (Zhang and Pierce, 2013). In tree fruit crops, the major aspects related to orchard architecture that affect the performance of a robotic harvesting system include plant spacing, canopy shape and canopy size. For example, variation in tree canopy training system substantially alters the percentage of visible and accessible fruit in the canopy. In United States, tree fruit orchards including apple and cherry canopies have been modernized rapidly over the past decade with a goal of reducing labor costs and increasing yield. The effort has also led to orchard training systems that are more compatible with mechanical or automated pruning, thinning and harvesting (Robinson, 2008). In a modern fruiting wall canopy architecture, the fruit visibility could reach up to 90% (Silwal et al., 2014). These studies have indicated that mechanization and automation of tree fruit harvesting cannot be effective if machine design is carried out in isolation of canopy design. So, a reconfiguration of the canopy structures should be considered in order to reach the degree of profitability expected when automating the harvesting task (Jimenez et al., 1999).

As mentioned above, numerous research and development efforts have been carried out in the past around the world in mechanizing or automating tree fruit harvesting operation. These techniques can broadly be classified in two categories: bulk harvesting and selective harvesting. Bulk harvesting methods can harvest all or most of fruit within target region at once. On the other hand, selective harvesting methods focuses on picking individual fruit based on their quality criteria such as color, ripeness, and size Some important features of these harvesting methods are discussed in the following sub-sections.
**Bulk Harvesting:** Bulk harvesting systems are designed for mass removal of fruit by exciting tree trunks, limbs or canopies. Such excitations could be achieved by inertial shaking devices, impacting devices or air blasting devices. Basic principle of impact-type harvesters is to accelerate each fruit so that an inertia force developed is greater than the bonding force between the fruit and the tree branch (Pacheco and Rehkugler, 1980; Markwardt et al., 1964). Impulse or impact shaking usually consists of application of one or more impacts on the trunk or the limbs of the tree. Fruit detachment results from the rapid acceleration exerted on the fruit. The major limitation of impact shakers was that the method achieved satisfactory results only on trees with a rigid structure, such as apricot and almond trees (Pacheco and Rehkugler, 1980). Continuous shaking is another major approach for bulk harvesting of fruit like apples, and cherries (Zhou et al., 2012; Peterson et al., 1999). Several researchers have studied fruit response to a sinusoidal excitation applied to tree limbs. These studies showed that the fruit detachment occurs after the fruit has developed an oscillatory motion of an adequate frequency and/or amplitude (Wang and Shellenberger, 1967; Diener et al., 1965).

Fruit removed in bulk by mechanical shaking method are accompanied with a catching surface to capture the detached fruit. Thus, this technique is also called shake-and-catch harvesting. Mechanical shaking techniques are being developed to deliver effective shaking energy to tree branches to achieve localized or targeted fruit removal (He et al., 2013; Zhou et al., 2013). On the other hand, catching system are being developed with specialized catching surfaces to minimize fruit damage while catching (Ortiz et al., 2011).
Selective Harvesting: Selective harvesting systems are designed to harvest fruit selectively based on predetermined criteria such as fruit color, shape, size and ripeness. To achieve such selective capabilities, the harvesters are equipped with specialized robotic arms and sensors. In 1968, Schertz and Brown (1968) proposed a concept of automated citrus harvester with a robotic arm for picking and machine vision to detect fruit and guide the arm. They used light reflectivity differences between leaves and fruit to detect fruit. Performance of their method was affected by non-uniform illumination and foliage density that limited the fruit visibility. Parrish and Goksel (1977) implementation a machine vision for apple detection. They used a Black-and-White camera along with a red optical filter to increase the contrast between red apples and green leaves. Since then, numerous research studies have been carried out for fruit detection and localization in various fruit crops including apples (Baeten et al., 2008), oranges (Slaughter and Harrell, 1989), peaches (Kurtulmus et al., 2014) and kiwi (Scarfe et al., 2009). Basically, the major steps involved in fruit detection for automated harvesting include image acquisition, preprocessing, noise filtering, segmentation, morphological operations, and feature extraction and interpretation (Gongal et al., 2015; Li et al., 2011).

1.2.2 Current Status of Sweet Cherry Harvesting

Washington State is the largest producer of fresh market sweet cherries in the United States producing more than 264,000 tons of sweet cherry annually (USDA, 2013). As mentioned before, no commercially adoptable harvester existed for mechanized or automated sweet cherry harvesting, which is a huge concern for the long term sustainability of cherry industry. Driven by the need for a sustainable method of harvesting, several studies have been conducted in the past to develop mechanical or automated cherry harvesters (Gatson et al., 1959; Norton et al., 1962;
One of the widely used harvesting technique involves the use of vibrational energy, which was shown by many researchers to be one of the most efficient ways of harvesting tree fruit crops, especially for harvesting smaller fruit growing in cluster like cherries (Chen et al., 2012; Du et al., 2011; Peterson and Wolford, 2001). In this harvesting method, the fruit removal is accomplished by delivering vibrational energy at appropriate frequency using trunk shakers, limb shakers, or canopy shakers (Ortiz and Torregrosa, 2013; Erdoğan et al., 2003; Peterson et al., 2003; Peterson and Wolford, 2003; Whitney et al., 1977), which detaches fruit at its abscission zone (Chen et al., 2012).

Peterson and Wolford (2001) developed a mechanical sweet cherry harvester with a Rapid Displacement Actuator (RDA) operated hydraulically that rapidly displaced tree branches for fruit removal. The removed fruit fell on a catching conveyor which then transported fruit into a collection bin. Chen et al. (2012) compared cherry harvesting system that uses impulsive excitation force with the system that uses series of vibratory excitations. It was reported that greater harvesting efficiencies could be achieved through intermittent vibrations of tree branches than by an impulse force, though more time is required for exciting branches with multiple vibrations. Successive evaluations of a prototype mechanical harvester revealed the difficulty for the operator to position the actuator due to limited viewing angle from the operator’s fixed seating position (Peterson et al., 2003; Larbi et al., 2015). A machine vision guided actuator could simplify the operation of such harvesters and could lead to fully autonomous cherry harvesting.

Previous studies showed that the mechanical shake-and-catch harvesting of cherries is a promising technique for mechanizing sweet cherry harvesting. Such mechanical harvesters could
be operated more efficiently and effectively with machine vision-based automated control. Machine vision and a decision support systems can be used to detect branches and determine shaking positions on those branches. A control system can then be used to position the shaking actuator in desired shaking locations. Avoidance of manual positioning with the use of a machine vision system will help develop a harvester with multiple layers of catching surfaces and with multiple actuators to improve fruit quality as well as harvest efficiency.

1.3 Goals and Objectives

The overall goal of this research is to develop a machine vision system for automated cherry harvesting using a branch shaking method. First step to achieve this goal is the automated detection of branches in cherry tree canopies. As the cherry trees are covered with leaves and clusters of cherries during the harvest season, a branch detection method needs to be capable of estimating branch location even when branches are heavily or sometimes completely occluded by dense foliage and fruit. In addition, it is necessary to estimate locations of cluster of cherries on each branch to make desirable harvesting decisions. Finally, based on the detected tree branches and the location of cherries, a number of shaking positions on each branch has to be determined. Hence, the following specific objectives are defined to achieve the overall goal of this work.

i. Detection of cherry tree branches in full foliage canopies using the morphological features of visible branch segments.

ii. Integration of visible branch sections and cherry clusters for detecting cherry tree branches in dense foliage canopies.
iii. Localization of shaking positions in cherry tree branches for automated sweet cherry harvesting.

1.4 Relevant Previous Research

A research group at the Center for Precision and Automated Agricultural Systems (CPAAS), Washington State University (WSU) has been developing mechanization technologies for sweet cherry harvesting. Over the past several years, the researchers have investigated various harvesting techniques including impact harvester, hydraulic shakers, and pneumatic and electric handheld shakers (Du et al., 2011; Chen et al., 2012; He et al., 2013; Zhou et al., 2013; Larbi et al., 2015). The difficulty in operating mechanical harvesters was experienced and reported by the researchers during field tests. These harvesters also required skilled operators to properly position mechanical shakers on tree branches. Therefore, fruit removal efficiency and harvesting time was highly dependent of operator’s skill level. The need of a machine vision system for automated positioning of shakers on tree branches was then realized. Additionally, the machine vision system will also enable us to develop multilayer shake-and-catch systems with multiple shakers. Such a system can harvest and catch cherries close to the branch and potentially minimize drop height and reduce fruit damage.

The automated detection of cherry tree branches and localization of shaking positions on tree branches will provide information to the robot control system to position the shaking mechanism at desired locations on tree branches for automated cherry harvesting. Therefore, this research will provide the knowledge and technology to bridge the gap between mechanical harvesting and automated harvesting of cherries.
1.5 Dissertation Organization

This dissertation was organized in six different chapters. In Chapter One, a brief background of the research and literature review was presented followed by the statement of main goal and objectives of this research.

Chapter Two will present some fundamental concepts on machine vision systems that are relevant to this research.

A method to detect cherry tree branches using the morphology of visible branch segments will be presented in Chapter Three.

Chapter Four will present a method to detect occluded tree branches based on the location of clusters of cherries in tree canopies.

Third objective of localizing shaking positions in cherry tree branches based on the distribution of cherry clusters in the canopy will be discussed in Chapter Five.

Finally, Chapter Six will present general conclusions drawn from this research and provides a few recommendations for the future research directions in this area.
References


Sarig, Y., (2005). Mechanized fruit harvesting- Site specific solutions. Information and Technology for Sustainable Fruit and Vegetable Production, FRUTIC 05, Montpellier, France.


A machine vision system, also termed as a computer vision system, can be defined as the ability of a machine or computer to visually perceive surroundings with the help of optical sensors. It has been widely used for automatic inspection, guidance, control, recognition, and navigation purposes in industrial as well as agricultural applications. As human beings are naturally gifted with the ability to perceive the world around us, it seems to be a trivial task to recognize the shape, size, color and distance of objects in our surroundings. To develop similar perception capabilities in a machine, it requires the understanding of visual systems and the development of optical sensors which can gather visual information from its surroundings. Researcher in computer vision have been developing sensing devices and mathematical techniques to detect various attributes of objects similar to human perception including shape, size, color, pattern/texture and location. In addition, machine vision systems have also been devised to sense object characteristics such as electromagnetic signals that are beyond human perception limits. An example of application of machine vision system is shown in Figure 2.1. The image of apple trees are acquired by using cameras in the apple orchard. Image acquisition step is followed by low, mid and high level processing to extract necessary information form the surrounding environment.
2.1 Digital Image Acquisition

An image can be defined as a two dimensional representation of a physical object or environment. Images are generated by the combination of an illumination source and the reflection, absorption and/or transmittance of energy by the elements of the scene being imaged (Figure 2.2). The signal received by the imaging sensor is analog and continuous, however such signals need to be converted to a digital form to be stored as a digital image. The process of conversion of analog signal into discrete data structure is called image digitization, which results in a matrix of real numbers representing an image (Gonzalez and Woods, 2002). Images captured by a monochromatic or Black-n-White (BW) camera can be represented by a two dimensional
matrix, $M \times N$, where $M$ and $N$ are rows and columns of the matrix. Each spatial coordinates of the image have a discrete value called intensity or gray level.

![Digital image acquisition process](image)

**Figure 2.2: Digital image acquisition process; energy emitted by illumination source is reflected by scene elements, which is captured by the imaging sensor and recorded as a digitized image (Gonzalez and Woods, 2002).**

### 2.2 Color Vision

A monochrome or gray image carries only intensity information. It is also called a black and white image in which lowest intensity is represented with black color and highest intensity with white color. Any other intensity values between the highest and lowest possible values are represented by different shades of gray. Color images, also called RGB image, are composed of three individual gray images for representing red (R), green (G) and blue (B) colors. A RGB image is a three dimensional matrix, $M \times N \times 3$, where $M$ and $N$ are number of row and column of a matrix. Therefore, RGB images can be viewed as the stack of three $M \times N$ dimensional
matrices. RGB images are defined based on an additive color model in which three primary colors (red, green and blue) can be added together in various ways to produce a broad arrays of colors. The addition of all three colors produces white color whereas the absence of all of them results in black color. Similarly, the addition of red and green yields yellow, green and blue yields cyan, and red and blue yields magenta respectively (Figure 2.3).

![RGB color model; additive colors: red and green produce yellow, green and blue produce cyan, red and blue produce magenta](image)

Color vision is a very important part of a machine vision system because it plays an important role in object recognition. Color is one of the commonly used features for distinguishing one object from another. For example, branch, leaf and cherry were all detected with the help of their color features.

RGB color model is the most widely used model in machine vision. But it is worth noting that there are several other color models besides RGB. For example, HSL and HSV are other two commonly used cylindrical coordinate color models (Figure 2.4). HSL stands for hue (H), saturation (S), and lightness (L), whereas, HSV stands for hue (H), saturation (S) and value (V). For these cylindrical models, the angle around the central vertical axis corresponds to hue, the distance from the axis corresponds to saturation and the distance along the axis corresponds to lightness. Hue is represented as primary red color starting at 0°, passing through green at
120° and blue at 240°. Saturation gives a measure of degree to which a pure color is diluted by white. Lightness or value describes the brightness of the color.

Figure 2.4: a) Hue, Saturation, and Lightness (HSL) color model and b) Hue, Saturation, and Value (HSV) color model (Source: Wikipedia)

2.3 Image Segmentation

Image segmentation is the process of subdividing an image into its constituent regions based on various features of the objects present in the image. The type of problem that is being addressed with the machine vision system dictates the level of subdivision required during the segmentation process. Image segmentation may stop when the object(s) of interest in an image has been isolated. Image segmentation was one of the important steps in this research which was used to segment out branch and cherry regions from leaves and background. The success of the machine vision system depends heavily on the accuracy of image segmentation process. Therefore, image segmentation is one of the most crucial tasks on a machine vision system. Some of the image segmentation techniques that are relevant to this research are discussed below.
2.3.1 Image Thresholding

Image thresholding is a simple image segmentation technique used to separate a group of pixels with similar gray levels from rest of the pixels constituting the image background. Thresholding is performed in a gray image which could be a monochrome image or one of the components from a color image. The assumption behind thresholding is that each object has different intensity level and by specifying a specific intensity range the object can be segmented out from its background. The process of selecting a suitable threshold value can be aided by analyzing the histogram of intensity levels (Figure 2.5). If there is a clear bimodal break in the distribution of pixels in the histogram (Figure 2.5, left), a single thresholding at $T$ may be sufficient to segment the objects of interest from the background. Sometimes, multiple thresholding will be required to segment the desired objects. For instance, in the histogram depicted by Figure 2.5 (right), there are three distinct groups of pixels over the range of intensity values. In such cases, depending on the object(s) of interest, thresholding may be applied at $T_1$ and $T_2$ to segment a group of pixels at the middle of the histogram from the rest of the pixels.

![Figure 2.5: Gray-level histograms depicting opportunities for image segmentation by single threshold (left) and multiple thresholds (right) (Gonzalez and Woods, 2002)](image-url)
2.3.2 Segmentation with Supervised Classification

Supervised classification could be a powerful tool for image segmentation when the object of interest does not have a distinct intensity or color features from that of the background. This method can be used to divide an image into multiple regions. Statistical classification methods, which use some prior information about the distribution of features of objects of interest, have been used widely in the past. One such supervised classification method for image segmentation is Bayesian Classification.

**Bayesian Classifier:** Bayesian classifier classifies an object into the class to which it is most likely to belong based on the observed features (Shapiro and Stockman, 2001). It is based on Bayes’ theorem which makes statistical interpretations from the prior knowledge and probability distribution of different class features. Let us assume a classification problem where objects have to be classified into two classes, \( w_1 \) and \( w_2 \). It is also assumed that the prior probability of both classes, \( P(w_1) \) and \( P(w_2) \) are known. For simplicity, let’s assume that only one feature \( x \) is being measured to make the decision about which class does the object belong to. Therefore, \( x \) would be a continuous random variable whose distribution depends on the class from which the object is drawn. This is denoted as \( p(x|w_i) \) and is defined as a class-conditional probability, where \( w_i \) represent the class \( (i=1 \text{ or } 2 \text{ in this case}) \). By applying Bayes Rule, the posterior probabilities for each classes can be calculated from the class-conditional pdf, \( p(x|w_i) \) and the priori probability \( P(w_i) \) (Duda et al., 1973) as follows.

\[
p(w_i|x) = \frac{p(x|w_i) \cdot P(w_i)}{P(x)} \tag{2.1}
\]

Where,
\[ p(x) = \sum_{i=1}^{n} p(x|w_i)P(w_i) \quad \text{(2.2)} \]

The product of class-conditional pdf and prior probability is the key term in determining the posterior probability because the term \( p(x) \) is simply a scale factor (Duda et al., 1973). Therefore, the decision function of Bayesian classifier can be written in the form:

\[ d_i(x) = p(x|w_i) * P(w_i) \quad \text{(2.3)} \]

Assuming the probability density functions are Gaussian, the \( n \)-dimensional Gaussian pdf has the form:

\[ p(x|w_i) = \frac{1}{(2\pi)^{n/2}|C_i|^{1/2}} \exp \left[ -\frac{1}{2} (x-m_i)^T C_i^{-1} (x-m_i) \right] \quad \text{(2.4)} \]

Where \( C_i \) and \( m_i \) are the covariance matrix and mean vector of the feature vector of class \( w_i \), and \( |C_i| \) is the determinant of \( C_i \).

As the logarithmic function is a monotonically increasing function, choosing the largest \( d_i(x) \) to classify the features is same as choosing the largest \( \ln[d_i(x)] \). Therefore, Eq. (2.3) can take the following form (Gonzalez, et al., 2009).

\[ d_i(x) = \ln P(w_i) - \frac{1}{2} \ln |C_i| - \frac{1}{2} [(x - m_i)^T C_i^{-1} (x - m_i)] \quad \text{(2.5)} \]

For a one dimensional feature vector, the probability distribution is two dimensional as shown in Figure 2.6. The point of intersection between probability distribution of two classes’ \( w_1 \) and \( w_2 \) gives the decision boundary. Therefore, if the measured feature has value on the left of the decision boundary, it is classified as class \( w_1 \), otherwise it is classified as class \( w_2 \).
2.4 Morphological Operations

Morphology or mathematical morphology in image processing refers to the shape or feature of the object in the image. Morphological image processing is a mathematical tool for performing various tasks including noise removal, void filling, region growing, object boundary extracting, skeletonizing and object thinning. Such operations are performed in images to remove imperfections caused during segmentation and other image processing operations. Morphological operations are implemented by passing a small binary image called structuring element over the image to be processed. The shape and size of the structuring element determines the specific impact morphological operation will cause on the images.

*Dilation* is one of the fundamental morphological operations that thickens or expands the objects in a binary image. The extent to which the object is expanded is determined by the shape and size of structuring element. *Erosion* is another fundamental operation that is used to shrink or
erode the objects in a binary image. Similar to dilation, the extent of erosion is controlled by the structuring element used. The dilation and erosion operations can be used in various combinations. Morphological *opening* is the process of erosion followed by dilation, which is generally applied to smooth the contours of an object by breaking and eliminating thin protrusions. Morphological *closing*, on the other hand, is the process of dilation followed by erosion. This operation is generally applied to fill small gaps and holes in the image as well as connecting broken or nearby objects by adding few layers of pixels at object boundaries. Region *filling* is another morphological operation commonly used to fill enclosed holes in images.

### 2.5 3D Vision

A 3D camera can capture a scene in three dimensions. There are two categories of 3D vision system, one is a range camera which produces a 2D depth image showing the distance of each pixels from a specific point, and another is a stereo camera which has a pair of 2D camera at slightly different viewpoints separated by a base distance. The stereo camera produces the distance information of the image pixels by stereo matching and triangulation method. In this research, a range camera based on time-of-flight (TOF) was used.

The working principle of TOF camera is shown in Figure 2.7. The object is actively illuminated with an incoherent light signal, which is intensity modulated by a cosine signal of frequency $f$. The reflected light is captured by the sensor and the phase difference between the emitted and reflected signal is used to calculate the depth of the object (Pankaj et al., 2013). Photonic Mixing Device (PMD) CamCube 3.0 (PMD Technologies, Seigen, Germany) is one example of a 3D
imaging system (Figure 2.8). It is a type of range camera which produces a 2D depth image providing the distance to each pixel from a specific point in the camera co-ordinate system.

**Figure 2.7: Working principle of time-of-flight-of-light (TOF) based camera (Shim and Lee, 2012)**

**Figure 2.8: Photonic-Mixer-Device (PMD) camera based on time-of-flight-of-light (TOF)**

### 2.6 Image Co-registration

Color vision provides color information of the objects which helps in color-based image processing for segmentation and extraction of various features and properties of the objects. On the other hand, a 3D vision system provides the depth information of the object which is
important for estimating the location and orientation of objects. In robotic applications, both the color and depth information is generally essential to detect and locate objects. Therefore, the images captured by 3D camera and color camera need to be mapped onto the same coordinate system such that the distance to objects detected in RGB images could be obtained from the corresponding 3D images. The RGB and 3D camera pair used in this research is shown in Figure 2.9.

![RGB camera (Bumblebee XB3, top) and 3D camera (PMD CamCube 3.0, bottom) pair used for mapping depth and color information](image)

**Figure 2.9:** RGB camera (Bumblebee XB3, top) and 3D camera (PMD CamCube 3.0, bottom) pair used for mapping depth and color information

The first step in mapping depth information on RGB image is the camera calibration. The camera calibration toolbox developed by Bougent (2013) can be used for calibrating each camera. From camera calibration the intrinsic camera parameters including focal length (F), principal point (C) and distortion coefficient (k) are determined. Using these intrinsic camera parameters, the stereo camera calibration can be carried out to determine extrinsic camera parameters, which were rotation matrix (R) and the translation matrix (T) of 3D camera with respect to the color camera.
For the 3D camera, the pinhole camera model can be written as equation 6 (Van den Bergh, M., and L. Van Gool, 2011).

\[
S. \begin{bmatrix}
u_{TOF} \\
v_{TOF}
\end{bmatrix} = \begin{bmatrix}
f_{x,TOF} & 0 & c_{x,TOF} \\
0 & f_{y,TOF} & c_{y,TOF} \\
0 & 0 & 1
\end{bmatrix} \cdot \begin{bmatrix} X \\
Y \\
Z
\end{bmatrix} \quad (2.6)
\]

Where \( u_{TOF} \) and \( v_{TOF} \) are the coordinates of a pixel in the image captured from the 3D camera, \( S \) is a scale factor, and \( X \), \( Y \), and \( Z \) are the coordinates of the corresponding 3D point.

The 3D coordinates can be computed as,

\[
\begin{bmatrix} X \\
Y \\
Z
\end{bmatrix} = Z_{TOF} \begin{bmatrix} f_{x,TOF}^{-1} & 0 & c_{x,TOF} \cdot f_{x,TOF}^{-1} \\
0 & f_{y,TOF}^{-1} & c_{y,TOF} \cdot f_{y,TOF}^{-1} \\
0 & 0 & 1
\end{bmatrix} \cdot \begin{bmatrix} u_{TOF} \\
v_{TOF}
\end{bmatrix} \quad (2.7)
\]

Where \( Z_{TOF} \) is the depth coordinates \((u_{TOF}, v_{TOF})\). \( X \), \( Y \) and \( Z \) values can be projected onto the coordinate system of the RGB camera as:

\[
\begin{bmatrix} X' \\
Y' \\
Z'
\end{bmatrix} = R \begin{bmatrix} X \\
Y \\
Z
\end{bmatrix} + T = Z_{TOF} \cdot R \begin{bmatrix} f_{x,TOF}^{-1} & 0 & c_{x,TOF} \cdot f_{x,TOF}^{-1} \\
0 & f_{y,TOF}^{-1} & c_{y,TOF} \cdot f_{y,TOF}^{-1} \\
0 & 0 & 1
\end{bmatrix} \cdot \begin{bmatrix} u_{TOF} \\
v_{TOF}
\end{bmatrix} + T \quad (2.8)
\]

Where \( X' \), \( Y' \) and \( Z' \) are the 3D coordinates with respect to the RGB camera.

The pinhole model for the RGB camera can be written as,

\[
S'. \begin{bmatrix} u_{RGB} \\
v_{RGB}
\end{bmatrix} = \begin{bmatrix}
f_{x,RGB} & 0 & c_{x,RGB} \\
0 & f_{y,RGB} & c_{y,RGB} \\
0 & 0 & 1
\end{bmatrix} \cdot \begin{bmatrix} X' \\
Y' \\
Z'
\end{bmatrix} \quad (2.9)
\]
Where \( u_{\text{RGB}} \) and \( v_{\text{RGB}} \) are the coordinates of the pixel in RGB image, and \( S' \) is a scale factor. Given that \( Z' \) is not zero, the coordinates in the RGB image can thus be computed using equation 10 and 11, which allows the projection of any pixel from depth image to the corresponding coordinate in RGB image.

\[
u_{\text{RGB}} = f_{x,\text{RGB}} \frac{x'}{Z'} + c_{x,\text{RGB}} \quad (2.10)
\]

\[
v_{\text{RGB}} = f_{y,\text{RGB}} \frac{y'}{Z'} + c_{y,\text{RGB}} \quad (2.11)
\]
References


CHAPTER THREE

DETECTION OF CHERRY TREE BRANCHES WITH FULL FOLIAGE IN PLANAR ARCHITECTURE FOR AUTOMATED SWEET CHERRY HARVESTING

(Reprinted from Biosystems Engineering, 2015, Amatya, S., M. Karkee, A. Gongal, Q. Zhang, and M. D. Whiting, Detection of cherry tree branches with full foliage in planar architecture for automated sweet-cherry harvesting, doi:10.1016/j.biosystemseng.2015.10.003, Copyright© (2015), with permission from Elsevier Ltd.)

Abstract

Fresh market sweet cherry harvesting is a labour-intensive operation that accounts for more than 50% of annual production costs. To minimise labour requirements for sweet cherry harvesting, mechanized harvesting technologies are being developed. These technologies utilise manually-placed limb actuators that apply vibrational energy to affect fruit release. Machine vision-based automated harvesting system have potential to further reduce harvest labour through improving efficiency by eliminating manual handling, positioning and operation of the harvester and/or harvesting mechanism. A machine-vision system was developed to segment and detect cherry tree branches with full foliage, when only intermittent segments of branches were visible. Firstly, an image segmentation method was developed to identify visible segments of the branches. Bayesian classifier was used to classify image pixels into four classes - branch, cherry, leaf and background. The algorithm achieved 89.6% accuracy in identifying branch pixels. The length and orientation of branch segments were then analysed to link individual
sections of the same branch together and to represent the branches with an equation. Linear and logarithmic model equations were fitted to the branch segments and the equation with minimum residual was selected as the best-fit model representing the corresponding branch. Branches detected with this algorithm were compared with manual counting. The method achieved a branch detection accuracy of 89.2% in a set of 141 test images acquired during full-foliage canopy. This study shows the potential of using a machine vision system for automating shake-and-catch cherry harvesting systems.

**Keywords:** cherry harvesting, image segmentation, Bayesian classifier, branch detection, vertical planar architecture, upright fruiting offshoots

### 3.1 Introduction

Washington State produced more than 264,000 tonnes of sweet cherry in 2012, which was 62% of total production of United States (USDA, 2013). Currently, all of these fruit are harvested manually, which is a highly labour intensive operation. Labour for harvesting constitutes more than 50% of total production costs (Seavert et al., 2008) and about 71% of the total human labour required for sweet cherry production (Employment Security Dept., 2013). As labour-related issues are becoming challenging due to increasing cost and decreasing availability (Gongal et al., 2015a; Fennimore and Doohan, 2008; Hertz and Zahniser, 2013), the interest in developing mechanical harvesting solutions has increased.

In recent years, sweet cherry growers in Washington State have been adopting new orchard training systems that are more compatible to mechanical harvesting (Peterson and Wolford, 2001; Long, 2010) than traditional systems. The upright fruiting offshoots (UFO) canopy
architecture is a modern, planar training system that consists of trees with a permanent horizontal limb from which multiple vertical limbs are grown (Whiting, 2009). The UFO system may be trained to a vertical or Y-trellised architecture which provides a compact fruiting wall. Such a system is amenable to mechanical or automated harvesting aided by a machine vision system for fruit and branch detection. Investigations on mechanical harvesting have shown the potential to improve harvest efficiency by adopting the UFO training system (Du et al., 2011; Chen et al., 2012). An economic study suggested that mechanically harvested cherry production systems will return more money to growers than the traditional system (Seavert and Whiting, 2011). The study was based on the results from a USDA mechanical harvester evaluated in early 2000’s (Peterson and Wolford, 2001). The actuator of this harvester was manually controlled by a joystick to position and engage a rapid displacement actuator (RDA) on a limb.

Evaluations of the prototype mechanical harvester revealed the difficulty for the operator to position the actuator due to limited viewing angle from the operator’s fixed seated position (Peterson et al., 2003). A subsequent study on mechanical harvesting of sweet cherry reported a significant effect of orchard characteristics and operator performance on the harvest rate (Larbi and Karkee, 2014). In addition, multi-layer catching surfaces located very close to the canopy may be essential to improve collection rate and reduce fruit damage rate during mechanical harvesting. However, this type of collection mechanism will critically limit the visibility and ability of an operator to localize branches for shaking. To address these issues, there is a need to develop an automated harvester using a machine-vision-based system for detecting shaking point in tree branches, and positioning the end-effector.
Systems and methods for mechanised cherry harvesting have been widely studied (Zhou et al., 2014; Larbi and Karkee, 2014; Du et al., 2012; Peterson et al., 2003; Peterson and Wolford, 2001; Halderson, 1966; Norton et al., 1962), yet only limited studies have investigated the potential for automating these harvesters. One study attempted to develop a cherry harvesting robot capable of picking individual cherries from tree canopies with the aid of 3D machine vision sensors (Tanigaki et al., 2008). The 3D sensors were attached to a robotic manipulator, which was able to pick the cherries that were visible to the sensors from a given viewpoint. However, cherries could be located all around the tree trunk. To minimise undetected cherries due to occlusions, the arm has to be moved to different viewpoints (Tanigaki et al., 2008). Fruit detection accuracy is critical for obtaining high harvesting efficiency because sweet cherry is characterised by many small fruit. Automated mechanical shakers may be more practical than robotic harvesting for crops like sweet cherry. One advantage of mechanical shaking method is that not every fruit needs detection as long as concentrated areas of fruit in branches are detected. For automatically harvesting cherries using mechanical shakers, a machine vision system needs to be capable of detecting and localising fruit as well as branches.

Studies have been reported in the past for detecting tree branches or similar structures in images. Detection of road network from aerial or satellite images using 2D image processing has been one of such studies (Hu et al., 2007; Laptev et al., 2000; Trinder and Wang, 1998). Morphological features has been widely used in these studies for detecting and joining road segments to extract the road network (Zhang et al., 1999; Valero et al., 2010). Lü et al. (2014) carried out fruit and branch identification for citrus harvesting robots in which branches were identified mainly for the purpose of obstacle avoidance while picking fruits. McFarlane et al.
(1997) used 2D image processing for pruning long wood grape vines. The branch images, taken during dormant season, were skeletonized for detecting branch centrelines and unconnected branch segments were joined using cubic curve equations. The method of joining branch segments with cubic curves was reported to be slow and only 60 pixels limit was applied for filling gaps in order to reduce processing time. He et al. (2012a) performed 3D reconstruction of Chinese hickory trees for the purpose of mechanical harvesting. However, the trees were imaged during dormant season without any leaves on them. In fact, most of the 3D branch reconstruction studies for orchard applications focused on trees in dormant season when branches are clearly visible (Gao and Lu, 2006; Karkee et al., 2014; Karkee and Adhikari, 2015).

The major contribution of this study was the detection of occluded branches of fruit trees during harvest season, which has not been studied widely in the past. Images captured during harvest season include dense foliage and fruit that occlude the complete view of branches and techniques used in dormant season branch detection could not be applied directly to this work. Therefore, the goal of this study is to detect cherry tree branches during full foliage season, which will provide a foundation for automating cherry harvesting operation using mechanical shaking of limbs. The specific objectives are to:

- Segment branch pixels in the cherry tree canopy images captured in the presence of leaves and fruit in orchard environment, and
- Detect individual branches using segmented branch regions and estimate their location and orientation in 2D images and assess the detection accuracy.
3.2 Materials and Methods

Images of cherry tree canopies were acquired at night using artificial illumination because the variability in natural illumination during day time could affect the robustness of the image processing algorithm. The images were then preprocessed to enhance contrast and reduce specular reflections. Branch pixel segmentation and noise filtering followed image preprocessing. A branch detection algorithm was then applied to the segmented images for detecting whole branches from partially visible sections. Performance of branch detection algorithm was assessed by comparing detection result with manual branch count. Detailed explanations of these methods are provided in the following sub-sections.

3.2.1 Test Site and Image Acquisition

The experimental data was collected in the Washington State University (WSU)’s experimental orchard at Prosser, WA, USA. Cherry trees of Skeena variety used for this study were trained in an Upright Fruiting Offshoots (UFO) architecture (Figure 3.1). The vertical offshoots were tied to 4 trellis wires at heights of 600 mm, 1080 mm, 1650 mm and 2200 mm from the ground respectively (Figure 3.2). The orchard had row spacing of 3000 mm with tree spacing of 1800 mm and canopy height of 3500 mm.
Imaging was carried out during night-time under controlled lighting conditions generated using white LED (Light Emitting Diode) lights, which produced an average illumination of 200 lux (+/-50 lux) in the imaging region. The camera used for image acquisition was Bumblebee ® XB3 (Point Grey Research Inc., B.C., Canada), which was a stereo vision device with three lenses with maximum resolution of 1280 × 960 pixels, aligned horizontally at 120 mm spacing. Each lens had a focal length of 6 mm with 43° Horizontal Field of View (HFOV). In this study, images were acquired at 512 × 394 pixels resolution. This research focused on detecting tree branches through segmentation of 2D images, therefore only those images obtained from central lens of the stereovision camera were used for the evaluation of branch detection method. Image processing and branch detection algorithm was implemented in MATLAB (Math Works Inc., Natick, MA, USA).
The camera was mounted on a pan-n-tilt system (FLIR Systems Inc., OR, USA) and the whole imaging system was mounted on a utility vehicle, Gator™ (Deere & Company, Moline, IL, USA) for field operation. The vehicle was positioned such that the distance between the camera and the tree row was approximately 1000 mm. For each tree, three images were taken at 15° increment of tilt angle of the pan-and-tilt system, so that the region between trellis wire 1 and trellis wire 4 was captured in the image (Figure 3.2). The lowest camera position captured images above wire 1 up to mid-section between wire 2 and 3. Mid camera position is the central part that included wire 2 and 3 whereas high camera position captured region including wire 3 and wire 4. There was approximately 50% overlap in images captured between each consecutive position. Some branches extended in the region above wire 4, but fruiting regions were mostly below wire 4. For validation purpose, images captured only from fruiting region (below wire 4) of the canopy were used. In total, images of 47 trees were used in this study.
Figure 3.2: Field image acquisition setup with artificial lighting. Camera mounted on a pan-n-tilt system took three images capturing region between wire 1 and wire 4.

3.2.2 Image Preprocessing

Preprocessing of images involved image enhancements and removing specular reflections. Image enhancement was performed by increasing the dynamic range of the image by stretching the contrast ratio of the images, which increased the overall contrast of the image. Image enhancement ensured that all images have the same dynamic range of image intensity values (Figure 3.3).
Figure 3.3: a) An image of cherry tree; and b) image after image enhancement by increasing contrast ratio

For removing specular reflection, firstly a mask of the regions with specular reflection was created by segmenting pixels with grey values above 250. For each of the red, green, and blue (RGB) channels, the mask region within images was filled by interpolating inwards from the mask boundary by solving the Laplace equation (Das et al., 2011). Figure 3.4 shows an example of removal of specular reflection from the image.
Figure 3.4: a) Original image with specular reflections on cherry centres and along medial axis of branches; and b) image after removal of specular reflections by inward interpolation method.

3.2.3 Pixel-Based Image Classification

The images of cherry trees consisted of branches, cherries, leaves and background (mostly dark pixels). Pixel-based image classification was carried out to classify pixels into different classes, which was necessary to segment branch pixel regions for implementing branch detection method. RGB values obtained after image preprocessing step were used as colour features. Firstly, the classification method was performed to classify the images into two classes; branches being the first class and everything else in the second class as background. To evaluate the potential for improving the accuracy of branch pixel segmentation, image classification was also carried out for 3 classes: branches, cherries and background; and 4 classes: branch, cherry, leaf and background. Training data set consisting of feature vectors for each class was created from 20 randomly selected images of cherry trees.

The average illumination over the imaging region was 200 lux with a variation of ±50 lux. The upper region had lower illumination compared lower regions. To account for this variability, the
training set was created including images from all regions of the canopy. The region of interest for each class was manually selected from the images ensuring that the selected region only included the pixels from the desired class, which were then stored as feature vectors. Manually selected sample regions contained a total of 345,554 pixels for branches, 206,643 pixels for cherries, 1,438,809 pixels for leaves and 1,707,734 pixels for background region. Since, leaves and background region occupied larger image area, larger sample regions were selected for those to incorporate variability within those classes. A feature vector created for each sample consisted of red, green and blue intensity values for each class.

3.2.3.1 Bayesian Classifier

There are variability in the pixels intensity values within a class due to the variability in colour signature of sample pixels as we all as variability in lighting condition. Bayesian Classifier was used for image classification as it accounts for such variability while making decision. This method classifies an object/pixel into the class to which it is most likely to belong based on the observed features (Shapiro and Stockman, 2001; Eq 1). The input parameters for Bayesian classifier is a multivariate feature vector \( x \) for all classes \( w_i \). The prior probability, \( P(w_i) \), which is also a key variable in the decision making, is defined as the percentage of feature vectors belonging to a class \( w_i \) with respect to total number of feature vectors. The class conditional probability density function (pdf), \( p(x|w_i) \) is the distribution of \( x \) given that it belongs to the class \( w_i \). By applying Bayes Rule, the posterior probabilities for each classes can be calculated from the class-conditional pdf, \( p(x|w_i) \) and the priori probability \( P(w_i) \) (Duda et al., 1973) as follows.

\[
p(w_i|x) = \frac{p(x|w_i) \cdot P(w_i)}{p(x)} \tag{3.1}
\]
Where,

\[ p(x) = \sum_{i=1}^{n} p(x|w_i)P(w_i) \] (3.2)

The product of class-conditional pdf and prior probability is the key term in determining the posterior probability because the term \( p(x) \) is simply a scale factor (Duda et al., 1973). Therefore, the decision function of Bayesian classifier can be written in the form:

\[ d_i(x) = p(x|w_i) * P(w_i) \] (3.3)

Assuming the probability density functions are Gaussian, the n-dimensional Gaussian pdf has the form:

\[ p(x|w_i) = \frac{1}{(2\pi)^{\frac{n}{2}} |C_i|^{\frac{1}{2}}} \exp\left[ -\frac{1}{2}(x-m_i)^T C_i^{-1} (x-m_i) \right] \] (3.4)

Where \( C_i \) and \( m_i \) are the covariance matrix and mean vector of the feature vector of class \( w_i \), and \( |C_i| \) is the determinant of \( C_i \).

As the logarithmic function is a monotonically increasing function, choosing the largest \( d_i(x) \) to classify the features is same as choosing the largest \( \ln[d_i(x)] \). Therefore, Eq. (3.3) takes the following form (Gonzalez, et al., 2002).

\[ d_i(x) = \ln P(w_i) - \frac{1}{2} \ln |C_i| - \frac{1}{2} [(x-m_i)^T C_i^{-1} (x-m_i)] \] (3.5)

### 3.2.4 Individual Branch Detection

After segmenting branch pixels using image classification method described above, a branch detection method was applied. For each segmented region, an ellipse with the central moments
equal to that of the region can be drawn as shown in Figure 3.5. For ellipses, two axes are defined; the major axis and the minor axis which is the axis perpendicular to the major (Figure 3.5).

Figure 3.5: Equivalent ellipse, major axis length and minor axis length of a segmented region (shaded region)

Steps involved in branch detection are shown in the flowchart in Figure 3.6. The segmented images consisting of branch pixels were post processed using morphological operations including dilation and erosion to fill holes and smooth the edges. Noise and other small regions were removed if the area was smaller than a specified threshold value.
Figure 3.6: Flowchart of the branch detection method; The input to the process is a segmented cherry tree image and the output is the number of branches in an image and equations representing those branches.
Most of the branches of the cherry trees used this study were approximately 50 to 60 mm thick (30–40 pixels, depending on the distance from camera). The length of the minor axis of segmented regions was used to detect branch sections with large thickness caused by erroneous segmentation of leaves as branches or by the segmentation of multiple branches together. A minor axis length threshold of 50 pixels was used allowing a tolerance of 10–20 pixels to accommodate slightly curved branches. Whenever the minor axis length of a segmented region was greater than the threshold, the region was skeletonised to its medial axis, which was carried out by iteratively deleting successive layers of pixels on the boundary until only the skeleton remains (Lam et al., 1992). Those points in the skeleton from which two or more branches emerge are called branching points. All the branching points of the skeleton were identified and the skeleton was broken into separate branch sections by removing those branching points (Figure 3.7). Individual sections were later connected to other appropriate sections to form longer branch segments.
Figure 3.7: Skeletonization and fragmentation of cherry branch segments; Medial axis skeleton was obtained and branching points were identified in the skeleton; Branches were then divided at branching points to form smaller, individual sections.

Another criterion used for detecting branch segments was the major axis length of segmented regions. As branches are elongated, the major axis of the branch is generally much longer than the minor axis and is oriented along its median axis. Those regions with major axis length greater than 50 pixels were detected as branch sections. The threshold value of 50 pixels has been chosen for minor axis length and major axis length allowing a tolerance for varying branch sizes in the images, which also accommodates minor fluctuations in sensor position with respect to tree canopies. The threshold value may have to be adjusted if the imaging system or distance to the objects differs significantly from the setup tested in this work.

With full foliage there is a high probability that only small sections of branches are visible, which would fail the major axis length criteria. Such smaller regions belonging to a single branch have to be detected and linked together to create the whole branch. Therefore, extraction of
geometrical features of all visible segments of branches was done and those features were then used to group individual sections of the same branch together.

For each branch segment, a search region was defined along its major axis with a specified lateral offset. The offset value used in this case was 25 pixels considering the thickness of the branch. However, the search region in the longitudinal direction (along the medial axis of the branch) was not restricted. That is because only about one-third of total tree height was visible in each images captured and it can be assumed that the branches extend well above and below the current field of view. In Figure 3.8, the dashed line represents the search area for the branch pixel region marked by the circle. Relative orientations were calculated for all other segmented regions within the search area. Whenever, the relative orientations of a given segment were within 20º of the orientation of the segment of interest, the segments were identified as parts of the same branch. All four sections shown in Figure 3.8 had similar orientation (i.e. relative orientation within 20º), therefore they were identified as parts of the same branch. The number of similarly oriented segments for the circled segment is 4 in this case. In order to be identified as a part of a branch, the segments must either have major axis length longer than 50 pixels or there needs to be more than two similarly oriented regions. Those regions which do not satisfy either criteria were not assigned to any branch. There could be some cases where major axis of the segments of the same branch may not be orientated in the same direction, particularly when a branch is curved. However, for the UFO tree architecture, the branches are generally upright and straight, and therefore the effect of this issue is minimal.
Figure 3.8: Orientations of branch segments represented by arrows; Two regions were assumed to be parts of the same branch when their relative orientations were within 20°.

All the segments that were detected as part of the same branch were labelled with a unique identifier. After all the partially visible branch sections were detected and labelled, an equation defining the trajectory of the branch was derived for each branch. Using the branch equation, the location of the branch segments in the occluded areas of the canopy could be estimated. Two types of models, a linear (for straight branches) and a logarithmic (for curved branches), were fitted through each branch to derive branch equations. The value of the minimum residual was used as a measure to identify the best-fit model equation,

Linear Model: \( Y = m \times X + C \)  \hspace{1cm} (3.6)

Logarithmic Model: \( Y = m \times \log(X) + C \)  \hspace{1cm} (3.7)

where, \( Y = \text{pixel position in rows} \)

\( X = \text{pixel position in columns} \)

\( m = \text{slope of equation} \)

\( C = \text{intercept} \)
Fitting a branch equation is useful for determining the location and orientation of branches with occluded segments as well as connecting individual segments of a single branch together. The curves generated by branch equations were plotted over cherry tree images for qualitative assessment of the fitness of the model to represent tree branches. Branch detection was considered successful when the curve representing the branch had same orientation and curvature as the actual branch. This was to ensure that if a point along this curve was picked for mechanical shaking, the shaker would come in contact and would be able to engage branch shaking at that location.

3.3 Results and Discussions

3.3.1 Pixel-Based Image Classification

The goal of image classification was to segment branch pixels in the cherry canopy images. To achieve this goal, Bayesian classification was implemented using grey levels in each of three colour bands (R, G and B, each varying from 0 to 255). Other colour features such as hue-saturation-value (HSV), and CIELAB were also evaluated during a preliminary test for image classification but RGB-based classification achieved comparatively better result. Figure 3.9 shows the decision boundaries in red-green, red-blue and blue-green planes developed using training dataset for two-, three- and four-class classification methods. The effect of including varying numbers of classes on decision boundary can be seen in Figure 3.9. Decision boundaries were presented in two dimensional space using two colour bands at a time. Inclusion of the ‘cherry’ class in the classification method altered the decision boundaries for both branch and background regions. Without the cherry class, 30.6% of the cherry pixels were classified into branch class resulting in a relatively lower branch classification accuracy. This change can be
observed by comparing decision boundary of two- and three-class classifiers (Figure 3.9). Similarly, the addition of the ‘leaf’ class further shrunk the branch and cherry decision boundaries. The leaves region extended to areas that would be under either branch or cherry classes with the three-class classifier. Thus by adding cherry and leaf classes, misclassification of these pixels into the branch class was reduced.

Figure 3.9: Decision boundaries for different classes of pixels in R-G, R-B, and B-G planes of a RGB colour space.
With two-class classification, in which branch pixels were segmented as foreground and non-branch pixels as background, the classifier achieved an accuracy of 75.6% (Table 3.1) and 72.4% (Table 3.2) for training set of 20 images and testing set of 80 images, respectively. The accuracy represents the percentage of actual branch pixels out of all the pixels classified as branch pixels. The high false positive (pixel identified as branch when actually not) result of 24.4% (Table 3.1) for two class classification was due to the misclassification of cherry pixels into branch class as evident from the decision boundaries (Figure 3.9). Another measure of the success of the algorithm was the number of false negatives. False negatives were such cases when a pixel from one class was misclassified into another class, primarily because of the overlap in colour features between two classes. From the red-green decision boundary for 3-class classification, it can be observed that the branch region is in between cherry and background regions. That means the mean colour signature of branch pixels lies between that of other two classes. Earlier with only two classes, cherries and background classes were merged together as a single class, which moved the mean and variance of the colour signature of the background class closer to that of the branch class. Since Bayesian classifier assumed the Gaussian distribution of the data, the actual cherry pixels were likely to be classified as branch pixels as seen in red-green decision boundary for 2-class classification. Having a separate cherry class reduced such likelihood and helped increase the classification accuracy.

With three-class classification, the accuracy increased to 87.0% for training and 84.3% for testing set, as previously misclassified cherry pixels were now correctly classified as cherry. The classification accuracy further improved for four-class classification with branch segmentation accuracy of 91.0% for training and 89.6% for testing set. The decision boundary (Figure 3.9)
showed that the branch region had been shrinking by the addition of each new class, which improved the accuracy of branch pixel classification by avoiding unwanted pixels being classified as branch pixels. Although the goal was to segment out only the branch pixels, including more classes in the classification process substantially improved the branch pixel classification accuracy (72% to 90% for testing dataset). Table 3.3 shows the confusion matrix for four class classification with the training data set. High consumer’s accuracy (fraction of correctly classified pixels with respect to all the pixels classified as that class; computed in rows) is desirable for branch detection as it ensures that there are lesser non-branch pixels being segmented as branch, thus minimising the error in estimating branch segment parameters such as minor axis length and orientation that were used for branch detection (next sub-section). The producer accuracy computed from the columns (Table 3.3) is the fraction of correctly classified pixels with respect to all the pixels belonging to that class.

**Table 3.1: Classification accuracy for two-, three-, and four-class classification with training dataset.**

<table>
<thead>
<tr>
<th>Class</th>
<th>Accuracy (%)</th>
<th>False Positives (%)</th>
<th>False Negatives (%)</th>
<th>Accuracy (%)</th>
<th>False Positives (%)</th>
<th>False Negatives (%)</th>
<th>Accuracy (%)</th>
<th>False Positives (%)</th>
<th>False Negatives (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Branch</td>
<td>75.6</td>
<td>24.4</td>
<td>13.4</td>
<td>87.0</td>
<td>13.0</td>
<td>15.0</td>
<td>91.0</td>
<td>9.0</td>
<td>13.2</td>
</tr>
<tr>
<td>Cherry</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>87.2</td>
<td>12.8</td>
<td>50.5</td>
<td>73.0</td>
<td>27.0</td>
<td>38.7</td>
</tr>
<tr>
<td>Leaf</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>90.9</td>
<td>9.1</td>
<td>6.8</td>
</tr>
<tr>
<td>Background</td>
<td>98.7</td>
<td>1.3</td>
<td>2.8</td>
<td>96.0</td>
<td>4.0</td>
<td>1.3</td>
<td>91.2</td>
<td>8.8</td>
<td>8.2</td>
</tr>
</tbody>
</table>
Table 3.2: Classification accuracy for two-, three-, and four-class classification with testing dataset.

<table>
<thead>
<tr>
<th>Class</th>
<th>Testing Set</th>
<th>2 classes</th>
<th>3 classes</th>
<th>4 classes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>False</td>
<td>False</td>
<td>False</td>
<td>False</td>
</tr>
<tr>
<td></td>
<td>Accuracy (%)</td>
<td>Positives</td>
<td>Negatives</td>
<td>Accuracy (%)</td>
</tr>
<tr>
<td>Branch</td>
<td>72.4</td>
<td>27.6</td>
<td>11.6</td>
<td>84.3</td>
</tr>
<tr>
<td>Cherry</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>79.7</td>
</tr>
<tr>
<td>Leaf</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Background</td>
<td>98.8</td>
<td>1.2</td>
<td>3.5</td>
<td>95.6</td>
</tr>
</tbody>
</table>

Table 3.3: Confusion matrix of testing set for four class classification

<table>
<thead>
<tr>
<th>Actual Class (Pixels)</th>
<th>Consumer Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Branch</td>
<td>Cherry</td>
</tr>
<tr>
<td>302731 (89.6%)</td>
<td>8480 (2.5%)</td>
</tr>
<tr>
<td>16155 (10.1%)</td>
<td>117252 (73.3%)</td>
</tr>
<tr>
<td>18546 (1.2%)</td>
<td>1221 (0.1%)</td>
</tr>
<tr>
<td>8122 (0.5%)</td>
<td>79690 (4.9%)</td>
</tr>
</tbody>
</table>

The sample result of Bayesian segmentation is shown in Figure 3.10 where image pixels are classified as either branch, cherry, leaf, or background region based on their colour features.
3.3.2 Branch Detection

Segmentation with four-class classification was used to detect and connect branch segments to identify individual branches in the cherry tree images. Linear or logarithmic equations representing the branches were used to plot the predicted branches and compare their location with location of actual branches in the images. Generally, branch shaking would occur in the region below 4th wire and therefore branch detection algorithm for automated harvesting was validated through analysing the images in that region only. The method detected most of the branches achieving an overall branch detection accuracy (true positives) of 89.2%. Out of a total
of 453 branches in 141 test images, 404 branches were detected correctly, missing only 49 branches (10.8% undetected branches or false negatives) (Table 3.4). There was an overall 16.1% of false positive branch detection, which was calculated as the percentage of falsely detected branches to the total number of branches (including both detected and undetected). Results showed that there were more falsely detected branches compared to undetected branches. In spite of preprocessing of images for specular reflection removal, the major reason of false positives was specular reflections from leaves, cherries and trellis wires used for training trees. Nearly all false positive branches (60 out of 73 false positive branches) were caused by false positive branch pixel segmentation. Which means that if segmentation error could be reduced to only a small percentage, false positive branch detection would be also be very minimal. Therefore, further improvement in image processing methods to minimise specular reflection could be useful. Alternately, specular reflections may be minimised while imaging using light diffusers and polarising filters. The undetected branches, on the other hand, mainly included those branches that were completely occluded by cherries and leaves. It was estimated that approximately 17 branches (out of 49 total branches) were missed by the branch detection method (i.e. false negative branches) because one or more sections of branches were missed during segmentation. In the future, information on the location of columns of cherries can be added to assist in locating branches bearing those cherries. Density of foliage may also provide useful information on branch segments occluded by leaves.
Table 3.4: Accuracy in detecting individual cherry branches

<table>
<thead>
<tr>
<th>Camera Position</th>
<th>Low</th>
<th>Medium</th>
<th>High</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Detected (%)</td>
<td>87.4%</td>
<td>92.1%</td>
<td>88.1%</td>
<td>89.2%</td>
</tr>
<tr>
<td>(no. of branches)</td>
<td>(139)</td>
<td>(139)</td>
<td>(126)</td>
<td>(404)</td>
</tr>
<tr>
<td>False Positives (%)</td>
<td>18.9%</td>
<td>15.2%</td>
<td>14.0%</td>
<td>16.1%</td>
</tr>
<tr>
<td>(no. of branches)</td>
<td>(30)</td>
<td>(23)</td>
<td>(20)</td>
<td>(73)</td>
</tr>
<tr>
<td>Undetected (%)</td>
<td>12.6%</td>
<td>7.9%</td>
<td>11.9%</td>
<td>10.8%</td>
</tr>
<tr>
<td>(no. of branches)</td>
<td>(20)</td>
<td>(12)</td>
<td>(17)</td>
<td>(49)</td>
</tr>
</tbody>
</table>

Figure 3.11 shows example results from the branch detection algorithm in which each detected branch has been plotted over the original image (Figure 3.11b and 3.11d). Corresponding original images are shown in Figure 3.11a and 3.11c. It was observed that occasionally, two or more segments of the same branch were detected as separate branches, particularly when the branch was curved. In an automated harvesting system, multiple detection of the same branch may lead to shaking at multiple locations along a branch. However, multiple shaking of a single branch will not likely affect the cherry removal efficiency. In fact, it is often desirable to shake branches at multiple locations to achieve high fruit removal efficiency (Zhou et al., 2014, Zhou et al., 2012; He et al., 2012b). Therefore, breaking of a single branch into multiple branches was not considered as a false positive while evaluating branch detection accuracy. It is noted that if a branch segment is relatively long, thin branches with a diameter of 3 pixels (approximately 5 mm) or more could be detected. However, if the visible segments of branches were really small with area less than 100 pixels, they were removed as noises, which was one of the reasons for false negatives. It is difficult to estimate the degree of occlusion that could be tolerated as it depends on various factors such as relative position, orientation, size and number of detected
branch segments. It was observed that most of the gaps of 250 pixels (400 mm) or less were filled successfully.

![Figure 3.11: Individual branch detection: original image (left); detected branches (right)](image)

It is noted that this work focused on detecting and locating branches in 2D images, though, localization of branches in 3D space is necessary for automated harvesting, which is the next step in the continuum of this work. Because a stereo vision system was used to acquire colour images, the same set of data can be used to evaluate the potential of using stereo-vision system for 3D localisation of branches. Stereo matching can be based on individual branches, which will
potentially improve accuracy and computational speed of the stereo-vision system. Alternative
way for 3D localisation could be the fusion of RGB images and depth information collected with
a 3D camera (Gongal et al., 2015b; de Bergh and Gool, 2011; Alenyà et al., 2011).

Having a capability to harvest cherries at night time also benefits grower to deliver higher quality
cherries to packaging houses because of the lower temperatures overnight compared to the day.
Lower temperatures help cherry to retain its desired quality measures such as firmness and sugar,
reduce decay as well as help maintain colour (Young and Kupferman, 2015). In addition, a
harvesting machine capable of working in both day and night-time is going to be more practical
to maximise the utilisation of expensive equipment. An over-the-row sensing system with a
tunnel structure has been developed in the past at Washington State University (Gongal et al.,
2015b; Gongal et al., 2014; Silwal et al., 2014), which has shown promise for both day and night
time operation. An automated harvesting machine with similar structure could be investigated in
the future.

3.4 Conclusion

This study was carried out to investigate the application of a machine vision system in
automating sweet cherry harvesting. Branch detection is the first step for automated harvesting
with mechanical shakers. The segmentation branch pixels in RGB images of cherry tree canopies
was done using a Bayesian classifier. The method achieved a branch pixel classification accuracy
of 89.6%. The morphological properties of the segmented branch sections were used to filter out
noises as well as to group together the segments of the same branch in a specified
neighbourhood. A curve fitting method was then used to fit an equation through detected branch
segments to connect them together. The overall accuracy in detecting individual branches was 89.2%.

The study showed promising results in detecting branches of sweet cherry trees during harvest season in the presence of full foliage and fruit. Information about the location of cherry fruit may be considered to improve branch detection accuracy. In addition, identification and localisation of cherries would be necessary future step to make decisions in terms of shaking location(s) within a cherry branch. Integration with 3D localisation, shaking end-effector, and path planning method(s) would then have to follow to develop and evaluate an automated harvesting prototype.

Acknowledgement

This work was supported in part by the USDA National Institute of Food and Agriculture (NIFA), Hatch project# 1005756 and project# 1001246 received from Washington State University Agricultural Research Center. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the view of the U.S. Department of Agriculture.
References


CHAPTER FOUR

INTEGRATION OF VISIBLE BRANCH SECTIONS AND CHERRY CLUSTERS FOR DETECTING CHERRY TREE BRANCHES IN DENSE FOLIAGE CANOPIES

Abstract

To minimize the demand of seasonal workforce in sweet cherry production, there is a need to develop automated harvesting systems. The first step for automating a shake-and-catch type harvesting system is to develop a machine vision system for detecting tree branches and localizing shaking point in those branches. In this study, an image processing algorithm was developed to detect branches of cherry trees using segmentation of branch and cherry pixels. First, partially visible branch segments within the tree canopies were connected using morphological features of the segments to form whole branches. Then, the positions of cherry clusters in the canopy were used as an indication to detect branch sections that were occluded by cherries and leaves. Different cherry clusters were grouped together based on their spatial location and distance between them. Branch equations were then defined through those cherry clusters using minimum residual criteria. In total, 93.8% branches were detected in a Y-trellis fruiting wall cherry orchard, with 55.0% of branches detected using only branch pixels and 38.8% additional branches detected using cherry clusters. The method resulted in a total of 12.4% of false positive detection. The results showed that branch detection accuracy can be substantially improved by integrating cherry location information with the location of segments of partially visible branches. This study has shown a potential for the application of machine
vision system to detect cherry tree branches in full foliage season, which is highly promising for the development of automated sweet cherry harvesting systems.

Keywords: Branch detection, shake-and-catch cherry harvesting, automated harvesting, Upright Fruiting Offshoots, Y-trellis canopy architecture

4.1 Introduction

Sweet cherry harvesting is a highly labor intensive operation which constitutes more than 50% of total production costs (Seavert et al., 2008). Currently sweet cherry harvesting is carried out manually by semi-skilled seasonal labor. Thousands of cherries growing in random spatial locations on individual tree canopies make commercial handpicking highly inefficient and costly (Li, 2011). With decreasing availability of agricultural workforce and increasing labor costs, developing automated solutions for cherry harvesting has been one of the most critical needs of the sweet cherry industry around the world.

There are several factors that hinder the development of automated harvesting solutions including technological, economical and horticultural factors. Tree architecture is one of the important factors affecting the harvesting efficiency (Ampatzidis and Whiting, 2013). In recent years, growers of Washington State and other part USA have been adopting more mechanization friendly architectures (Peterson and Wolford, 2001; Long, 2010). The Upright Fruiting Offshoots (UFO) canopy architecture is one of the examples of a modern, planar training system that consists of trees with a permanent horizontal limb from which multiple vertical limbs are grown (Whiting, 2009). The UFO system may be trained to a vertical or Y-trellised architecture and provides a compact fruiting wall canopies. Investigations on mechanical harvesting have shown
the potential to improve harvest efficiency by adopting the UFO training system (Du et al., 2011; Chen et al., 2012). This architecture is also amenable to automated harvesting aided by a machine vision system for fruit and branch detection.

Mechanization of harvesting operation for various type of tree fruit and nuts have been widely investigated in the past with commercial success for various crops including nuts and fruit destined for processing market. One of the most widespread harvesting techniques investigated and commercialized in the past is mechanical shaking (He et al., 2012), which has been investigated for several kinds of tree fruit crops including pistachio (Polat et al., 2006), apricot (Erdogan et al., 2003), olive (Blanco-Roldán et al., 2009), apple (Peterson et al., 1999) and mango (Parameswarakumar and Gupta, 1991). A mechanical harvester developed and comprehensively evaluated by USDA in early 2000’s (Peterson and Wolford, 2001) was based on engaging a rapid displacement actuator (RDA) on a limb using a manual controller. Evaluations of the efficiency of this prototype mechanical harvester revealed the difficulty for the operators to position the actuator due to limited viewing angle from the operator’s fixed seated position (Peterson et al., 2003). A subsequent study of the harvester reported a significant effect of orchard characteristics and operator performance on the harvesting speed (Larbi and Karkee, 2014). In addition, multi-layer catching surfaces located very close to the canopy may be essential to improve collection rate and reduce fruit damage rate during mechanical harvesting (observation based on ongoing work at Washington State University). However, this type of collection mechanism will critically limit the visibility and ability of an operator to localize branches for shaking. To address this issue, there is a need to develop an automated harvester
using a machine-vision-based system for detecting branches, identifying shaking points and positioning the end-effector.

Various investigations have been conducted in the past to detect and reconstruct branches and trunks of fruit trees (Tabb, 2009; Wang and Zhang, 2013; Karkee et al., 2014; Karkee and Adhikari, 2015). However, most of these studies have focused on detecting branches in dormant season with potential application in pruning and crop-load management. Amatya et al. (2015) developed a method for detecting branches using partially visible branch sections in full foliage canopy, which reported an accuracy of 89.2% in a vertical architecture with relatively lighter canopy density. With higher density of foliage and cherry clusters, the branch visibility may decrease drastically, limiting the accuracy of branch detection method using visible branch segments. However, as cherries grow along the branches, location of cherry clusters can be useful in estimating location of branches that are hidden by cherries and leaves. The specific objective of this study is to integrate the sections of partially visible branches and location of clusters of cherries to detect cherry tree branches in the presence of dense foliage during harvest season. The result can be used to guide the actuating mechanism of a harvester to tree branches for automated cherry harvesting.

4.2 Materials and Methods

The branch detection method primarily includes image acquisition in an experimental cherry block, followed by segmentation of branch and cherry pixels. Then, fully or partially visible branches were detected using segmented branch regions. In addition, segmented cherry clusters were used to detect occluded branches. Final branch count in each images were determined
combining results from branch pixel-based detection and cherry pixel-based detection methods. The performance of branch detection algorithm was assessed by comparing the final result with results from manual branch counting in images. More details in these steps are provided in the following paragraphs.

4.2.1 Image Acquisition

Image acquisition for this research was carried out in the WSU experimental orchard in Prosser, WA. The cherry trees were trained to Y-trellis architecture with fruiting limbs oriented about 55° from the horizontal plane. The test orchard had a tree spacing of 4.3 m X 1.7 m and tree height of approximately 3.5 m. To avoid the variability in natural illumination, image acquisition was carried out during the night with controlled lighting conditions provided by LED lights (Trilliant® 36 Light Emitting Doide Grote, Madison, Indiana). A stereo vision device (Bumblebee ® XB3, Point Grey Research Inc., B.C., Canada) was used in this study to acquire images. Bumblebee camera included three lenses with maximum resolution of 1280x960 pixels, aligned horizontally at 12 cm spacing. Each lens had a focal length of 6 mm with 43° Horizontal Field of View (HFOV). The camera was mounted about 0.5 m above ground on a utility vehicle, GatorTM (Deere & Company) such that the camera is pointed vertically upwards (Figure 4.1).
4.2.2 Image Classification

Bayesian Classifier was used to classify image pixels into four classes, namely: branches, cherries, leaves and background. Bayesian Classifier is a pattern classification technique based on probability theory, which uses the knowledge of probability distribution to make classification decisions with least expected error rate (Shapiro and Stockman, 2001). This technique is based on Bayes Rule, which estimates the posterior probability of classifying a pixel into a class based on the similarity of features as well as knowledge about the priori probability (prior) of occurrence of a particular class (Duda et al., 1973). The input parameters for Bayesian
classifier is a multivariate feature vector $x$ for all classes $w_i$. The Bayesian classifier has the decision function of the form:

$$d_i(x) = \ln P(w_i) - \frac{1}{2} \ln |C_i| - \frac{1}{2} [(x - m_i)^T C_i^{-1} (x - m_i)]$$

(4.1)

Where $P(w_i)$ is the prior probability, which is defined as the percentage of feature vectors belonging to a class $w_i$ with respect to total number of feature vectors. $C_i$ and $m_i$ are the covariance matrix and mean vector of the feature vector of class $w_i$, and $|C_i|$ is the determinant of $C_i$ (Gonzalez, et al., 2002).

The Bayesian classifier was trained using 20 cherry tree images. Training images were manually segmented into branch, cherry, leaf and background classes. Red, green and blue intensity values were extracted for each class and used as the feature vector for training the classifier. Detailed explanation on image classification method used in this work can be found in Amatya et al. (2015), which has reported a classification accuracy of 91.0%, 73.0%, 90.9% and 91.2% respectively for branch, cherry, leaf and background classes. The mean and covariance of the feature vectors estimated by Amatya et al. (2015) were used as Bayesian classification parameters for image classification in this study. The pixel-level classification accuracy for cherry tree images used in this study was evaluated with a test data set of 50 randomly selected images.

4.2.3 Branch Detection using Branch Pixels

A method developed by Amatya et al. (2015) for detecting tree branches using visible branch segments was applied to detect cherry tree branches. The detection method used the morphological features of visible branch segments including orientation, length and thickness to
group branch segments together. Individual branch segments were identified as the part of the same branch when two or more segments were oriented in same direction. Additionally, all branch segments longer than 8 cm were also detected as individual branches. Every branch detected by the algorithm was represented by an equation that approximated the medial axis of the branch. Linear model and logarithmic models were fitted to the branch pixels and the model with least residual value was selected as the best fitted branch equation. For detailed description of branch detection method using visible branch segments, please refer to Amatya et al. (2015).

4.2.4 Combining Branch and Cherry Regions for Improved Branch Trajectory

The previous method of branch detection could yield broken branches when the orientation of different visible sections of same branch differed by more than 20°, which could lead to the detection of the same branch multiple times. Figure 4.2(a) shows an example of branch regions obtained after image segmentation of cherry tree images. By implementing the orientation and length-based branch detection method discussed in previous subsection, two branch sections oriented in different directions were detected as shown in Figure 4.2(b). Integration of branch and cherry regions in the neighborhood of previously detected branches was carried out to improve the accuracy of branch detection method. Starting from the longest detected branch section, the branch section was extended by approximately 15 cm (100 pixels) on either side along the corresponding branch trajectory defined by the model equation. The region within 3 cm (20 pixels) on the lateral direction of the branch trajectory was considered as the neighborhood zone. If any branch or cherry region fell within the neighborhood zone, those branch and cherry regions were included to be the parts of the same branch and the branch equation was updated. The branch was then extended along the new branch trajectory to identify
if any other branch or cherry regions are in its neighborhood. This process continued until no further addition took place. Figure 4.2(c) shows the end result of this branch merging method, which combined the previously unassigned branch regions to form a longer branch. In addition, cherry regions in the neighborhood of branch trajectory (vertically striped; Figure 4.3c) were also merged with detected branch region to form longer branch sections.

Figure 4.2: a) Segmented branch regions (gray); b) Detection of branch; striped regions represented detected branch segments; gray regions represent branch regions not detected as a part of any branch; c) improved branch detection by combining neighborhood branch regions to form a single branch
4.2.5 Cherry Pixel-Based Branch Detection

Detection of branches based on branch pixels was successful in canopy regions with low foliage density where branches were fully or partially visible. In the canopy regions with high foliage and fruit density, branches were heavily occluded (some even completely occluded) leading to a substantial proportion of undetected branches. In addition, branch visibility was critically low when the tree branches were small in diameter with dense foliage and cherry clusters around them. In such conditions, the method for detecting branches based on only branch pixels led to a low branch detection accuracy. Detection of cherry clusters provided useful information in
detecting branches that are occluded by foliage or cherry clusters. Cherry clusters are located in a close proximity to the branch bearing them and completely occluded some branches. In a UFO cherry orchard used in this study, a number of cherry clusters aligned closely in a vertical direction indicated the presence of a branch along those clusters. When multiple cherry clusters were identified in such a fashion, a branch was assumed to be located there and a curve was fitted passing through those clusters.

Before using cherry clusters for detecting occluded branches, cherry cluster associated with visible branches were removed to avoid multiple detection of previously detected branches. This was achieved by assigning all cherry regions near the detected branches as the part of the same branch and excluding such regions in the following steps of branch detection. The segmented cherry pixels and the branch equations identified using branch sections were used as the inputs in this step of detecting branches with cherry pixel information. For each of the previously identified branches, a neighborhood search region was defined along the branch section (defined by the equation) with a specified lateral offset proportional to the width of the branch. The neighborhood search method is similar to the one carried out for updating branch equations as described in previous sub-section, except that a much larger lateral offset (minimum 50 pixels; approx. 8 cm) was used in this case. Any cherry region found in the neighborhood was marked with the respective branch identifier but the branch equation was not updated this time. After all cherry cluster in the neighborhood of detected branches were labelled, the remaining unmarked cherry regions were used in the next step for detecting new branches using only spatial location of cherry clusters. The major steps involved in detecting cherry tree branches using location of
cherry clusters is shown in the flowchart below (Figure 4.4) and described in the following paragraphs.

Figure 4.4: Flowchart showing major steps in branch detection using branch and cherry pixel regions
Remaining cherry regions, which were not associated with any previously detected branches, were used to predict new branches. Unlike branch segments, orientation of cherry clusters cannot be used as an indicator of branch orientation because the clusters are irregular in shape. Therefore, to detect branch orientation in this case, a region growing approach was used. Starting at the bottom of the image and gradually moving upwards, cherry clusters potentially belonging to the same branch were identified. First, the bottom-most cherry region was located in the segmented image and was labeled with a unique identifier as a parent region. A neighborhood was defined covering 200 pixels above and 15 pixels below the topmost point of the region and 75 pixels in the lateral direction around the centroid of the region (Figure 4.5). These parameters were selected using trial and error method for the given tree architecture, camera resolution and imaging distance. If any other cherry region was found in that neighborhood, those were labeled with the same label as the parent region and the centroid of newly added regions were located. Denoting the centroid of parent region as \((x_1, y_1)\) and the centroid of new region/regions as \((x_2, y_2)\), the total offset of centroid in x-axis was equal to \(x_2 - x_1\). New neighborhood search region was then defined above the new region with an offset value of \((x_2 - x_1) / 2\) in the x-axis (Figure 4.6). This centroid offset value was used to adjust the trajectory of search region according to the potential orientation of the branch to which these cherry clusters belong. This process of neighborhood search was repeated until no new regions were found.
Figure 4.5: A neighborhood for identifying cluster of cherries around a given cluster for predicting a branch passing through them.

Figure 4.6: Progression of the region growing method with a neighborhood search around a cherry region; a) first step from bottom-most region with no offset in x-axis; b) second step with an x-axis offset of (x2-x1)/2; and c) third step with x-axis offset of (x3-x2)/2.
When no more cherry regions were found in the vertical direction, the next unassigned cherry region at the bottom was located as a new parent region. Then the same process of neighborhood search was repeated to add new cherry clusters to the new parent region. After all cherry regions were grouped with individual parent regions, the overall area of each group was estimated. If the cluster group had the pixel area less than 1000 pixels (roughly three cherries or less), those groups were rejected from the potential list of branches. For each remaining cherry groups, a linear and a logarithmic models were fitted. One of the two equations with minimum residual was then selected as the best equation defining each of those branches. Two sets of branch equations determined in this and previous sub-sections were then combined as the final result of branch detection method.

4.3 Results and Discussions

4.3.1 Pixel Classification

Image pixels were classified as one of the four classes: branches, cherries, leaves or background. The mean and covariance of feature vector for each class that were used for image classification (Amatya et al., 2015) are shown in Table 4.1. Performance of the classification method was evaluated with a set of 50 test images (Table 4.2). Consumers’ accuracy, which is defined as the fraction of correctly classified pixels with respect to all pixels classified as that class, for branch and cherry pixel were found to be 81.4% and 95.5% respectively. Relatively lower consumers’ accuracy was achieved for the branch class because of misclassified pixels to this class from cherry and leaf classes. Cherry pixels, on the other hand, was classified with high consumers’ accuracy. However, producers’ accuracy, which is defined as the fraction of correctly classified pixels with respect to all the pixels belonging to that class, was estimated to be only 60.0% for
cherry pixels. It was also observed that about 35.2% of cherry pixels were misclassified as background. The main reason was the dark color of cherries, which was similar to the background color. Mostly the cherry pixels from the areas with low illumination were classified as background class. For the application in branch detection, it was desirable to have higher consumers’ accuracy to ensure that smaller percentage of misclassified pixels are present in a segmented region, which can be removed using morphological operations and other noise filtering techniques.

Table 4.1: Mean feature vector for branch, cherry, leaf and background classes

<table>
<thead>
<tr>
<th>Class</th>
<th>Red (R)</th>
<th>Green (G)</th>
<th>Blue (B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Branch</td>
<td>128</td>
<td>145</td>
<td>155</td>
</tr>
<tr>
<td>Cherry</td>
<td>39</td>
<td>304</td>
<td>38</td>
</tr>
<tr>
<td>Leaf</td>
<td>76</td>
<td>130</td>
<td>94</td>
</tr>
<tr>
<td>Background</td>
<td>12</td>
<td>21</td>
<td>15</td>
</tr>
</tbody>
</table>

Table 4.2: Confusion matrix depicting the performance of pixel classification method

<table>
<thead>
<tr>
<th>Actual (Pixels)</th>
<th>Consumers</th>
<th>Actual (Pixels)</th>
<th>Consumers</th>
<th>Actual (Pixels)</th>
<th>Consumers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Branch</td>
<td>39953</td>
<td>3498</td>
<td>5060</td>
<td>597</td>
<td>81.4%</td>
</tr>
<tr>
<td>Cherry</td>
<td>1189</td>
<td>48066</td>
<td>135</td>
<td>928</td>
<td>95.5%</td>
</tr>
<tr>
<td>Leaves</td>
<td>2563</td>
<td>355</td>
<td>445715</td>
<td>22263</td>
<td>94.7%</td>
</tr>
<tr>
<td>Background</td>
<td>1070</td>
<td>28177</td>
<td>22696</td>
<td>443893</td>
<td>89.5%</td>
</tr>
<tr>
<td>Producers Accuracy</td>
<td>89.2%</td>
<td>60.0%</td>
<td>94.9%</td>
<td>94.9%</td>
<td></td>
</tr>
</tbody>
</table>
4.3.2 Tree Branch Detection

A total of 138 images were evaluated for this study containing a total of 453 cherry tree branches. An example of overall branch detection process is shown in Figure 4.7. Figure 4.7(a) shows the original image of cherry tree and Figure 4.7(b) shows the identified branches and their respective equations. The lines with square markers in the figure represent branches identified using branch pixels whereas the lines with circular markers represent branches identified using cherry pixels.

It was found that 55.0% of branches were identified using only the branch pixels (Table 4.3). The detection accuracy with only branch pixel information was low because a substantial number of branches were occluded by the foliage and cherry clusters. Remaining cherry regions were used, after excluding previously identified branch and cherry pixels, to help predict occluded branches. An additional 38.8% of previously invisible branches were detected, leading
to the total branch detection accuracy of 93.8%. The rate of false positives in branch detection was 4.6% and 7.7% while using branch pixels and cherry pixels respectively. False positives were defined when there was no actual branch existed in the image or when the actual location and/or orientation of the branch were different than the predicted.

It was observed that the main source of false positive branch detection was specular reflections from leaves and low spacing between branches in some poorly trained trees leading to falsely segmented branch regions. Improvement in image classification method to decrease false positive pixels during segmentation may help reduce false detection of branches. In addition, a well-trained orchard with consistent spacing between branches could be helpful in reducing false positives. About 6.2% of branches were still unidentified after integrated branch detection method. False negative cases were primarily caused by poor lighting on the upper part of the tree canopy as well as thin branches with low fruit density. Improving the illumination condition during image acquisition without inducing specular reflections could be helpful in reducing unidentified branches and improving overall performance of the branch detection method.

Table 4.3: Results of branch detection method based on branch pixels and cherry pixels

<table>
<thead>
<tr>
<th></th>
<th>Based on Branch Pixels</th>
<th>Based on Remaining Cherry Pixels</th>
<th>Overall</th>
<th>Unidentified</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Detected</td>
<td>False Positives</td>
<td>Detected</td>
<td>False Positives</td>
<td>Detected</td>
</tr>
<tr>
<td>Number of Branches</td>
<td>249</td>
<td>21</td>
<td>176</td>
<td>35</td>
<td>425</td>
</tr>
<tr>
<td>Percentage (%)</td>
<td>55.0</td>
<td>4.6</td>
<td>38.8</td>
<td>7.7</td>
<td>93.8</td>
</tr>
</tbody>
</table>
4.4 Conclusion

Automated sweet cherry harvesting can be achieved with the aid of machine vision system for locating tree branches and guiding harvesting actuators to the desired locations for shaking. The main challenge for machine vision system is detecting branches in tree canopies with full foliage when a lot of branches are occluded. In this study, a branch detection algorithm was developed to identify cherry tree branches in canopies with full foliage during the harvest season. The study was conducted in a cherry orchard trained in Y-trellis architecture. Bayesian classification method was used to classify image pixels into branches, cherries, leaves and background achieving classification accuracies of 81.4% and 95.5% for branches and cherries respectively. Segmented branch and cherry regions were used for the branch detection. First, sections of partially visible branches were used to reconstruct entire branches based on the geometrical properties of segmented branch regions. In addition, cherry clusters in the images were used to locate the branches that were totally occluded by the foliage and fruit itself. Based on branch pixels only, 55.0% of cherry branches were identified, whereas the detection accuracy was improved by 38.8% using cherry clusters leading to the overall branch detection accuracy of 93.8%. Integration of branch and cherry pixel regions resulted in improved overall branch detection accuracy.

This study showed a huge potential for branch detection in full foliage, which can lead to an automated cherry harvesting system in the future. Further improvement in the performance of this method might be possible by improving pixel classification technique, uniform illumination in imaging region, and proper training of trees ensuring well-spaced vertical branches. The next step in automating shake-and-catch harvesting of sweet cherries is to locate the shaking positions
in the detected cherry tree branches. The position of tree branches has been detected along with the position of cherries in the detected branches. Therefore, based on the location of cherry clusters in the canopy, location of suitable shaking points for removing cherries has to be estimated for guiding mechanical shakers to the desired positions.

Acknowledgements

This work was supported in part by the USDA National Institute of Food and Agriculture (NIFA), Hatch project #1005756 and project #1001246 received from Washington State University Agricultural Research Center. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the view of the U.S. Department of Agriculture.
References


CHAPTER FIVE

A METHOD FOR LOCALIZING SHAKING POSITIONS IN
CHERRY TREE BRANCHES FOR AUTOMATED
SWEET CHERRY HARVESTING

Abstract

Automation in cherry harvesting is essential to reduce the cost of cherry production and reduce the demand of seasonal labor for cherry picking. Mechanical shaking of tree branches is one of the widely used techniques for harvesting small tree fruit like cherries. Different methods of detecting branches and cherries in full foliage canopies of cherry trees have been developed previously. The next step for automating the cherry harvesting process is the localization of shaking positions in the detected tree branches for mechanical shaking. In this study, a method of locating shaking positions for automated cherry harvesting was developed based on branch and cherry pixel locations determined using RGB images and 3D camera images. First, branch and cherry regions were located in 2D RGB images. Depth information provided by 3D camera was than mapped on to the RGB images using standard stereo calibration method. The overall root mean square error in estimating distance of desired shaking points was found to be 0.064m. Cherry trees trained in two different canopy architectures, Y-trellis and vertical trellis systems, were used in this study. Harvesting test was carried out by shaking tree branches at the locations selected by the algorithm. For Y-trellis system, maximum fruit removal efficiency of 92.9% was achieved using up to five shaking events per branch. Whereas, maximum fruit removal efficiency for vertical trellis system was found to be 86.6% with up to 4 shaking events per branch.
However, it was found that only three shakings per branch would be enough to achieve 92.3% and 86.4% fruit removal in Y and vertical trellis systems respectively.

**Keywords:** Cherry harvesting, shaking locations, 3D mapping, fruit removal, Y-trellis system, vertical-trellis system

5.1 Introduction

Cost and availability of human labor is one of the major concerns for sweet cherry growers. As cherries are characterized by clusters of small fruit in random spatial orientations on tree canopies, hand picking is very labor intensive. It has been reported that more than 50% of total cost of cherry production is incurred by harvesting operation. Though the labor-saving technology for sweet cherry harvesting is a critical need for the industry (Halderson, 1966), no commercially adoptable sweet cherry harvesters have been available to the growers so far. The advancement in mechanical sweet cherry harvesting would help cherry industry to be more competitive and sustainable in the long term (Seavert and Whiting, 2011).

Over the past several decades, many studies have been conducted to develop mechanical and automated solutions for harvesting tree fruit (Zhao et al., 2011; Baeten et al., 2007; Tanigaki et al., 2008; Ceres et al., 1988). One of the widely used harvesting technique is using vibrational energy, which has been shown to be the most efficient way of harvesting tree fruit, especially crops such as berries and cherries with smaller fruit growing in clusters (Chen et al., 2012; Du et al., 2011; Peterson and Wolford, 2001). In this harvesting method, fruit removal is accomplished by delivering vibrational energy at appropriate frequency using trunk shakers, limb shakers, or canopy shakers (Ortiz and Torregrosa, 2013; Erdoğan et al., 2003; Peterson et al., 2003; Peterson
and Wolford, 2003; Whitney et al., 1977) which generally detaches fruit at its abscission zone (Chen et al., 2012).

Peterson et al. (1999) proposed a fully automated bulk apple harvester with imaging system to guide a robotic arm. Automatic image segmentation was also attempted using a color camera. It was reported that detection of apple was successful but the detection of branches was considerably difficult. Therefore, picking shaking positions on the branches was done manually by clicking at the desired point on the tree image, which would provide the two dimensional coordinates (x, y) of the shaking location and the depth was determined physically by allowing the actuator to move until a limit switch was pressed when it came in contact with the branch. The development of a robust image processing system that can detect branches and a method to automatically locate shaking positions in the tree branches is the desirable next step towards fully automated tree fruit harvesting with a mechanical shaker.

Amatya et al. (2015) developed a method of detecting cherry tree branches covered with leaves using morphological properties of visible branch segments and reported a detection accuracy of 89% in vertical planar architecture. Amatya and Karkee (2015) improved the branch detection method by using location of cherry clusters as a clue to detect heavily or completely occluded branches in high foliage density orchards trained in Upright Fruiting Offshoots (UFO) architecture. The improved method reported a branch detection accuracy of 93% in a UFO orchard architecture with Y-trellis system. These studies showed a promise for a machine vision system to automate shake-and-catch cherry harvesting system. The next step in the automation of cherry harvesting system is to automatically determine and locate the shaking positions in the detected tree branches for shaking off cherries.
There have been many studies in improving mechanical harvesting technology for sweet cherries to increase harvesting efficiency and quality of harvested fruit while keeping the harvest-induced damage to the tree at a minimum level. Some studies have indicated that higher fruit removal efficiency can be obtained by prolonged shaking at multiple shaking positions (Zhou et al., 2013; Blanco-Roldán et al., 2009). Zhou et al. (2013) showed that fruit removal efficiency is affected by shaking frequency and duration, and reported 81% fruit removal with four intermittent shaking of 5 seconds each at a frequency of 18 Hz. Other studies have indicated that excitation position on the tree branches also influence the transfer of energy along the branch and consequently affects the fruit removal efficiency (Zhou et al., 2014; Adrain et al., 1965). Zhou et al. (2014) evaluated fruit removal efficiency with excitation at different positions by dividing Y-trellis cherry trees into 4 excitation zones. The results indicated that maximum fruit removal was achieved for lowest shaking position followed by the highest shaking position using handheld shakers. It was also reported that up to 97% fruit removal efficiency could be achieved if shaking was done at both top and bottom excitation zones.

Peterson and Wolford (2001) developed a mechanical cherry harvester with a branch impacting mechanism and catching conveyor system, which transported harvested cherries to a collection bin. It was reported that a potential for harvesting up to 85-92% of sweet cherries but also resulted in damage to the tree and fruit with multiple impacts during harvesting. Other major drawback was the difficulty in seeing the branches and subsequent difficulty in positioning the impacting mechanism on tree branches. Chen et al. (2012) also reported that the limited visibility for operator was challenging for accurately aiming the shaking mechanism to targeted branch locations. Larbi et al. (2015) modified Peterson and Wolford (2001)’s harvester by replacing the
impacting mechanism with a continuous shaking mechanism in order to reduce the damage caused by the impact. In addition, the positioning of the shaking mechanism was achieved by a remote controller to improve the harvester’s operability. With a remote controller in hand, the operator had more flexibility to move around to get a proper view and target the shaking mechanism accurately. With the added flexibility by using a remote controller, the efficiency of hitting a target branch with impactor was 93% and average time required for such maneuvering was 19.9 seconds per position. The results indicated that the positioning of shaking mechanism in target branches was still problematic because the presence of catching conveyor under the target branch would limit the operator from going too close to have a better view. On the other hand, the operator’s skill level affects the harvest rate and positioning time. Larbi and Karkee (2014) showed that operator experience level could result in 8.3% variation in shaker positioning time and up to 16.1% variation in fruit removal rate.

The difficulty in positioning the shaking mechanism on target branches is due to the limited visibility of the branches. A machine vision system can aid in detecting tree branches in dense canopies and position the shakers in target locations. This automated method will not only address the visibility problems but also reduce positioning time taken by operators. The detection of tree branches has been carried out in previous studies with satisfactory accuracy. However for positioning of shaking mechanisms on proper shaking locations, the shaking positions need to be identified in detected branches. Therefore, this research has focused on developing a method of automatically selecting shaking positions in tree branches for effective cherry harvesting. The shaking positions are selected considering the amount of cherries available for harvest and their
location in the canopy. The effectiveness of selected shaking locations is estimated in the field with harvesting experiments.

5.2 Materials and Methods

5.2.1 Test Orchard

The study was conducted in the Washington State University experimental orchard in Prosser, Washington. Cherry trees trained in two different architectures were used in this research. First cherry block was of Skeena variety trained in Upright Fruiting Offshoots (UFO) system with vertical limbs (Figure 5.1a). The vertical offshoots were tied to 4 trellis wires at 0.6 m, 1.08 m, 1.65 m, and 2.2 m above the ground. The orchard had row spacing of 3 m with tree spacing of 1.8 m, and approximate canopy height of 3.5 m. The second cherry block was of Selah variety trained in UFO Y-trellis system with fruiting limbs oriented about 55° to the horizontal surface (Figure 5.1b). The Y-trellis orchard had a tree spacing of 4.3 m X 1.7 m and tree height of approximately 3.5 m.
Figure 5.1: a) Upright Fruiting Offshoots (UFO) vertical planar architecture; and b) UFO Y-trellis architecture

5.2.2 Image Acquisition

Image acquisition was carried out at night with artificial illumination to avoid variation in natural lighting condition. LED lights (Trilliant ® 36 Light Emitting Doide Grote, Madison, Indiana) were used for illuminating the imaging region. Color images of cherry trees were acquired using a Bumblebee ® XB3 (Point Grey Research Inc., B.C., Canada) camera, which was a stereo-vision device with three lenses. The RGB images captured by the central camera with a focal length of 6 mm, Horizontal Field of View (HFOV) of 43° and resolution of 1280x960 pixels were used for detecting branches and cherries. A time-of-flight (ToF) based 3D camera (PMD CamCube 3.0, PMD Technologies, Siegen, Germany) was used to capture the depth information.
5.2.3 Co-registration of Depth and RGB Images

RGB images were analyzed for detecting and reconstructing tree branches using various color and geometric features of tree branches and cherry clusters. A time-of-flight-based 3D camera was used along with a RGB camera to obtain depth information of detected branches. The mapping of depth information from 3D camera images onto RGB images was carried out using standard stereo camera calibration algorithm form MATLAB camera calibration toolbox (Bougent, 2013). The resolution of 3D camera was 200x200 pixels whereas the resolution of RGB camera was 1280x960 pixels. To match the resolutions, 3D images were first up-sampled by linear interpolation to 1280x960 pixels.

The calibration algorithm provided the intrinsic camera parameters including focal length (F), principal point (C) and distortion coefficient (k) (skew coefficient for radial and tangential distortions). Using these intrinsic parameters, a stereo calibration was performed to determine extrinsic camera parameters, which included rotation (R) and translation (T) vectors of the 3D camera with respect to the RGB camera (Gongal et al., 2015; Bougent, 2013). For the co-registration of 3D images with RGB images, point PN at image plane of 3D camera was inverse mapped to point P in the 3D space (Figure 5.2). Point P was then transformed to 3D reference frame of the color camera using rigid motion transformation and finally projected on to the color image plane PR. For more details on co-registration process used in this work, please refer to Gongal et al. (2015).
5.2.4 Branch Detection and Reconstruction

Detection of cherry tree branches was carried out in color images acquired in the test orchards using methods developed by Amatya et al. (2015) and Amatya and Karkee (2015) (Figure 5.3). Cherry tree branches were covered by dense foliage, which limited the visibility of branches in tree canopies. A branch detection method was implemented to estimate the location of branches in canopies using partially visible branch sections. The intermittently visible branch sections were segmented from the images to evaluate their morphological features including orientation, length and thickness. Using these features as clues, the branch sections belonging to a branch were identified, and a model equation was fitted to reconstruct branches. Detailed explanation of
this method can be found in Amatya et al. (2015). This method detected most of the branches when at least a few sections were exposed to the camera. However, some branches were completely occluded by leaves or cherry clusters. To detect such branches, cherry location-based branch detection method (Amatya and Karkee, 2015) was implemented. For this method, cherry regions were segmented from images and a series of cherry clusters growing close to each other in a particular direction were identified as potential fruit clusters grown in a specific branch. Model equations were then fitted to represent the branches occluded by those cherry clusters. Detailed description on this method can be found in Amatya and Karkee (2015). The final output of branch detection methods was the mathematical equations representing all branches in the images.
5.2.5 Determining Shaking Locations in Tree Branches

The shaking position(s) in tree branches affects the fruit removal efficiency during shake-and-catch harvesting. Previous studies on optimizing shaking frequency, duration and position (Zhou et al., 2014; Zhou et al, 2013) provided a guidance for estimating initial shaking locations. Studies have shown that multiple shaking location would be essential for maximum fruit removal.
(Zhou et al., 2014; Zhou et al., 2013). Based on the fruit distribution and dynamics of tree branches, shaking on both upper canopy region and lower canopy region have potential to yield maximum fruit removal efficiency. Therefore, the initial localization of the shaking position in this study was carried out in three specific canopy regions (referred to as primary shaking positions). If any cherries were left on the tree after shaking on those initial locations, new shaking locations were determined based on the location of remaining cherries (referred to as secondary shaking positions). Details on localization of shaking position are provided in the following sub-sections.

**Determining Primary Shaking Positions**

The input for locating shaking position were branch and cherry regions obtained after image segmentation, and branch equations provided by the branch detection methods. First, the canopy region captured by an image was divided vertically into three zones: zone 1 (top), zone 2 (middle), and zone 3 (bottom) (Figure 5.4). For a given branch, it was determined if there were any cherry clusters assigned to the branch section in each canopy zone. A shaking position in the given zone of tree branches was determined only when cherries were present in that zone so that any unnecessary branch shaking can be avoided. Figure 5.4 shows an example image, in which branch 1 has cherries in zone 1 only. In such a case, there was no need to find shaking position in zone 2 and zone 3. Similarly, branch 2 had cherries in zone 3 requiring shaking at zone 3 only. Branch 3 has no or only inconsiderable number of cherries in it and therefore shaking would not be required in any zones.
Once cherries were located in all three shaking zones, a shaking position was determined to harvest those cherries. To reduce the possible damage to cherries, it was desirable to avoid direct contact of the shaker with cherries. Therefore, first priority was given to visible sections of the branch within the shaking zone on which desired cherry clusters are located. But the branch section may or may not always be visible in every situation due to the possibility of occlusion by fruit or foliage. The decision making process for different scenarios that may occur in this process (Figure 5.5) is described in the following paragraphs.
Figure 5.5: Flow chart of the shaking point localization process in tree canopies
Scenario 1: Branch segment is visible and satisfies branch equation

In this case, cherry clusters were detected in the given shaking zone and a part of the branch segment to which the cherries belong to was visible (white region in Figure 5.6b). Black dotted line (Figure 5.6b) is the trajectory of the branch represented by the branch equation, which was determined through branch segment detection and branch reconstruction method described in sec. 5.2.4. As the branch equation passes through some part of the segmented branch region, the pixel coordinates of overlapping regions will satisfy the branch equation. After the coordinates of branch region satisfying the branch equation were identified, the median location (which divided all satisfying coordinates into two halves, each half being in one side of the median location) was selected as the shaking position for the corresponding branch in the given canopy zone.

Figure 5.6: a) An example zone of a cherry tree branch showing clusters of cherries and a visible branch segment; b) Segmented branch region (white region), branch trajectory defined by the corresponding branch equation (black-dotted line). The median position of overlapping co-ordinates between branch region and branch trajectory was selected as the shaking position.
Scenario 2: Branch Sections Visible but does not Satisfy Branch Equation

In this scenario, branch segments were visible in the desired zone along with cherries, however branch equation did not pass through the visible branch section (Figure 5.7). In such cases, the branch section nearest to the branch trajectory was identified. Then the centroid of that branch region was identified as the desired shaking position for the given zone.

![Figure 5.7: a) An example zone of a cherry tree branch showing clusters of cherries and a visible branch segment; b) A scenario when branch equation does not pass through branch region; red dotted line represent the path defined by branch equation, segmented branch and cherry regions are shown as gray and blue striped regions](image)

Scenario 3: Branch Sections not Visible

Figure 5.8 depicts one example with completely occluded branch section. Branch equation, in this case, passed through the cluster of cherries, which occluded the branch section. In such
cases, the shaking position was selected allowing 2-3 cm (15 pixels) buffer zone for engaging the shaking mechanism below the largest cherry cluster.

Figure 5.8: A part of a cherry tree canopy where the branch section has been completely occluded by fruit and leaves; dashed line represents the trajectory defined by the branch equation and the circle below the largest cherry cluster represents the shaking position.

Fruit Harvesting using Primary Shaking Positions

As discussed before, shaking positions for cherry harvesting were located on tree branches at different canopy heights. Tree canopy was divided in three zones and shaking locations were estimated for each zone containing cherry clusters. Next step was to perform harvesting test by shaking branches on shaking locations determined by the algorithm. Harvesting was done by shaking tree branches at the locations identified by the algorithm using a handheld shaker. The handheld shaker was operated at a frequency range of 14 Hz-18Hz. Intermittent shaking of branches was continues up to 4 times with approximately 5 seconds of excitation every time.
This shaking frequency and harvesting pattern was implemented on the basis on previous studies on mechanical cherry harvesting (Zhou et al., 2013), which recommended such method for most efficient fruit removal. At most, three shaking locations were predicted per branch. However, previous studies (Zhou et al., 2014) had indicated that shaking in top and bottom zones would result in maximum fruit removal. Therefore, assuming that two shaking locations might be enough, the top and bottom shaking locations were harvested first. Mid zone was shaken only if there were any cherries remaining in that zone after shaking at top and bottom zones. The harvest test was performed by shaking tree branches in the following sequence:

i) Shaking at the identified location in zone 1 (top zone)

ii) Shaking at the identified location in zone 3 (bottom zone)

iii) Shaking at identified location in zone 2 (mid zone), only if there were cherries remaining after previous shakings

In ideal cases, up to three primary shaking positions would be sufficient to harvest all cherries in a tree branch. But there were cases when there could be inefficient energy transfer from shaker to branches, most likely due to thin and long branches. Such limitation may result in cherries not harvested completely with shaking at primary shaking positions. On the other hand, some branches could be un-detection in the first round due to occlusions. In either case, secondary shaking locations need to be determined to harvest remaining cherries. Hence, a second image of the tree canopy was taken and branch and cherry detection was carried out after completing shaking at the primary shaking positions. If more cherries were detected in the canopy, new shaking positions (secondary positions) were determined based on the location of detected
cherries. The method of determining secondary shaking positions is described in the following section.

_Harvesting with Secondary Shaking Positions_

It was assumed that the remaining cherries in tree branches did not detach during the primary round of shaking because of ineffective energy transfer. Therefore to maximize the energy transfer to remaining cherries, secondary shaking positions were determined such that shaking would occur close to the remaining cherry clusters. The input to this method was also the segmented images containing branch and cherry regions as well as equations of the detected branches. However, in this case instead of dividing the image into different zones, the whole field of view was used. First, the intersection between the trajectory defined by the branch equation and the cherry regions was determined (Figure 5.9). The pixel distance from each coordinate of the branch path to the nearest cherry region was evaluated. The distance to the nearest cherry region is shown as bars in the cherry distance profile (Figure 5.9c) along the branch path. For all those branch coordinates that overlap with cherry region, the distance to the nearest cherry region is zero. In Figure 5.9(b), point A lies within the cherry region and the distance to the nearest cherry cluster is zero as shown in the cherry distance profile. To avoid direct impact on the cherries while shaking, the shaking position should be picked outside the cherry region. On the other hand, it should be picked close to the cherry cluster to transfer maximum energy to cherries being harvested. Therefore, allowing a buffer zone of 2-3 cm (15 pixels), all coordinates that are 15 pixels away from the nearest cherry region were selected as potential shaking positions.
In Figure 5.10, potential shaking positions are shown as P1, P2, P3 and P4, which were at a distance of 15 pixels from the nearest cherry region. For every such positions, the impact region was defined as a square area of 500x500 pixels with the shaking position as its center. The area of cherry region inside each impact region was evaluated. The shaking position with the maximum cherry area inside the impact region was picked as the first shaking position. If more cherry regions remained in the tree, another shaking position, which has the maximum impact area for remaining cherries, was picked for shaking. This process was continued until all cherry regions were under an impact region that has already been shaken.
After determining secondary shaking locating, further harvesting tests were conducted. When multiple shaking locations were identified in a single branch, the harvesting was carries out from the top to bottom. The weight of cherries removed in each shaking was recorded.

5.3 Results and Discussions

5.3.1 Mapping 3D depth information onto RGB image

The distance to shaking positions estimated with mapping of depth information onto RGB images was compared with manual distance measurements using a laser distance measurement unit (DLR 130K, Bosch, Stuttgart, Germany). A total of 138 shaking positions were used for evaluating the performance of distance estimation method. As discussed in the method section,
the shaking positions were located on the visible branch sections when possible. However, branch sections were not always visible and the shaking position could be located on occluded regions as well. In such cases, distances to the shaking positions were estimated by linearly interpolating the depth information available for visible branch or cherry regions around the occluded shaking position. The errors in distance estimation for sample locations is shown in Figure 5.11. Overall 32% of shaking positions were located on the visible branch regions. The mean absolute error for distance estimation over the shaking positions within visible branch regions was 3.4 cm with a standard deviation of 3.4 cm (Table 5.1). Root mean square error for such cases was 4.8 cm. The distance to occluded branch regions was estimated with a mean absolute error of 5.2 cm with a standard deviation of 4.8 cm and the root mean square error of 7.1 cm. The estimation error was higher for occluded branch region as expected because the depth information was not directly available for those regions. Overall, the root mean square error for distance estimation was 6.4 m. Figure 6 depicts errors in estimating the distance to all branch sections with solid line representing mean error, and dashed lines representing upper and lower deviation lines. It is desirable to minimize the distance error so that the shaking mechanism could be guided precisely to the desired shaking position. Further improvement in distance estimation accuracy may be possible by improving the stereo camera calibration process. The low resolution of 3D camera is another factor that could affect the calibration accuracy.

Along with accurate distance estimation, the efficient harvesting also depends on the mechanical design of the shaker. Some level of errors in distance measurement can be compensated by designing shakers with a degree of tolerance. For example, the shaker could be built with wide
contact surface or a V-shaped hook with wider opening. The shakers used in previous studies have wide shaking head with around 30 cm (1 feet) width to allow proper contact with tree branches even if there is some offset in positioning the shaker. Considering such designs of the shaking head, the error of 6.4 cm (RMSE) could still be acceptable for successful cherry harvesting. Force sensors and limit switches can also be used to detect when the shakers come in contact with the branch.

![Figure 5.11: Error in estimating distance to branch sections through mapping of 3D information onto RGB images. Solid line represents mean whereas shaded band represent standard deviation region.](image)

**Table 5.1: Error in estimating distance to shaking positions from the 3D camera.** Distance was estimated by mapping 3D depth information onto color images.

<table>
<thead>
<tr>
<th>Branch</th>
<th>Mean Error (m)</th>
<th>Mean Abs Error (m)</th>
<th>RMSE (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Branch</td>
<td>0.015 ± 0.046</td>
<td>0.034 ± 0.034</td>
<td>0.048</td>
</tr>
<tr>
<td>Non-Branch</td>
<td>0.029 ± 0.065</td>
<td>0.052 ± 0.048</td>
<td>0.071</td>
</tr>
<tr>
<td>Overall</td>
<td>0.025 ± 0.060</td>
<td>0.047 ± 0.064</td>
<td>0.064</td>
</tr>
</tbody>
</table>
5.3.2 Fruit Removal Efficiency

The decision making algorithm for selecting shaking positions on cherry tree branches generated multiple shaking positions on each branch based on the distribution of cherries on individual branches. The number of shaking positions required per branch varied according to the tree training system and the fruit load on the target branches. The progression of fruit removal efficiency over increasing number of shaking positions has been analyzed for two training systems investigated. It was observed that the shaking at the first position always yielded the maximum fruit removal efficiency (percentage), which was 65.1% in Y-trellis canopy system (Figure 5.12a) and 67.7% in the vertical trellis system (Figure 5.13a). Shaking at the second shaking position on a branch also helped to remove significant amount of fruit from the branches bringing the cumulative fruit removal efficiency to 85.5% and 83.3% for Y-trellis system (Figure 5.12b) and vertical trellis system (Figure 5.13b) respectively. Shaking at the third shaking position removed only 6.5% fruit in Y-trellis and 3.0% in vertical trellis system to bring the overall fruit removal to 92.3% and 86.4% respectively.

As discussed in the methods section, cherry location-based shaking decision was made if there were more cherries on the tree after third shaking. In Y-trellis system fourth and fifth shaking yielded 0.54% and 0.02% fruit bringing the maximum fruit removal efficiency to 92.9% (Figure 5.12). Whereas, in vertical-trellis system the fourth shaking yielded 0.2% fruit achieving the maximum fruit removal of 86.7% (Figure 5.13). Shaking was not continued beyond fourth shaking for vertical trellis system either because no more cherries were detected or no fruit were removed by fourth shaking. It was observed that there was no significant improvement in overall fruit removal beyond three shakings per branch in most of the branches. Based on this result, an
upper limit of three shaking events per branch could be imposed for efficient harvesting. It was also observed that Y-trellis system with average branch diameter of 5 cm has overall higher fruit removal accuracy of 92.9%. Vertical trellis system had lower fruit removal efficiency of 86.6%, primarily because of undetected branches and cherries. Tree canopies in this architecture were as well trained and pruned to maintained the two dimensional structure as in the Y-trellis architecture, which might be the primary reason for lower harvesting accuracy. The results and field observations indicated that maintaining a good two dimensional fruiting wall structure and a certain spacing between branches may help improve overall harvest efficiency.

Figure 5.12: a) Fruit removal efficiency (percentage) for shaking at each consecutive position per branch on Y-trellis canopy system; and b) Cumulative fruit removal efficiency after each consecutive shaking on Y-trellis canopy system
Results also showed that 29.1% of branches in Y-trellis architecture required only one shaking to remove all removable cherries from the branch. For vertical trellis system, 47.4% branches were harvested completely with a single shaking per branch (Figure 5.14). It was observed that effectiveness of a shaking event in getting fruit detached was affected by branch diameter. Branches in the vertical trellis system had larger diameter compared to those in Y-trellis system, which potentially facilitated the effective transfer of kinetic energy along the branch causing larger fruit removal percentage per shaking. In Y-trellis system, two shaking events were required for 39.8% of branches whereas 31.5% of branches required three or more shaking for maximum fruit removal. For vertical trellis system, 36.8% of branches required two shaking event and 15.8% branches required three or more shakings respectively (Figure 5.14). These
results showed that the number of shaking events per branch should be decided dynamically depending on the load and distribution of cherries that is present in the canopy at the beginning and after each shaking event.

![Bar graphs showing the number of shaking positions required for maximum fruit removal in Y-trellis system (a) and in vertical trellis system (b).](image)

*Figure 5.14: Number of shaking positions required for maximum fruit removal in Y-trellis system (a) and in vertical trellis system (b).*

Previous research studies in mechanical cherry harvesting have generated a lot of knowledge on effective shaking method for harvesting cherries. This research focused on developing a method for automatically determining shaking positions on tree branches based on canopy and cherry locations. To develop an automated harvester, this information on the location of shaking positions will be essential for guiding shaking mechanism on target branches accurately and efficiently. The results of harvesting tests have also indicated that multiple shakings would be required on each branch for maximum fruit removal. With machine vision-guided shakers, the automated harvester could be equipped with multiple independent shakers at different canopy position with separate catching surface for each shaker. Such systems has a potential to improve
the efficiency and speed of the harvester as well as improve the quality of harvested fruit by reducing the drop height.

5.4 Conclusion

For automated sweet cherry harvesting with shake-and-catch system, the machine is required to make decisions on number of shaking positions and their 3D location in the canopy. This research focused on developing a method for determining shaking positions in cherry tree branches detected by a machine vision system. The localization of shaking position on tree branches included; i) Determination of shaking position in each branch to harvest cherries using RGB images; and ii) Estimation of distance to shaking position from the camera location by mapping depth information collected by a 3D camera onto the RGB images. The root mean square error (RMSE) on estimating distance to shaking positions was found to be 6.4 cm. It was also observed that the distance estimation was more accurate with RMSE of 4.8 cm when the shaking position was selected over the visible branch regions compared to the positions in occluded regions of the canopy where distance was estimated using a linear interpolation method.

The mechanical shaking of tree branches at the shaking positions determined by the algorithm was carried out in Y-trellis and vertical trellis canopies. First shaking event removed the largest amount of fruit from tree branches regardless of the tree architecture. The maximum fruit removal achieved with shaking at multiple positions was 92.9% for Y-trellis system, which required as many as 5 shaking per branch in some cases. For vertical trellis system, the maximum fruit removal efficiency was 86.6%, which took up to 4 shaking per branch. However,
it was found that three shaking position per branch would be enough for harvesting most of the cherries that could be removed by branch shaking in most of the branches. First shaking event in the vertical trellis removed more fruit (47.4%) compared to that in Y-trellis system (29.1%). The results indicated that relatively larger diameter (of vertical system) might play a role in increasing the effectiveness of energy transfer along the branch and therefore more efficient fruit removal. Overall fruit removal in vertical trellis system was lower than Y-trellis system because of undetected branches and cherries. Maintaining a good two dimensional fruiting wall structure and a certain spacing between branches may help improve branch and cherry detection accuracy and the overall harvest efficiency.

**Acknowledgement**

This work was supported in part by the USDA National Institute of Food and Agriculture (NIFA), Hatch project #1005756 and project #1001246 received from Washington State University Agricultural Research Center. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the view of the U.S. Department of Agriculture.
References


CHAPTER SIX

CONCLUSIONS AND RECOMMENDATIONS

6.1 General Conclusions

Fresh market sweet cherry harvesting is a highly labor intensive and costly operation. The development of automated systems for sweet cherry harvesting is essential for the long-term sustainability of the industry that is dealing with rising labor cost and increasingly uncertain labor availability. This research focused on detecting cherry tree branches in full foliage canopies and locating shaking positions for automated shake-and-catch harvesting using machine vision. Mechanical shaking of tree branches is a widely investigated technique for harvesting small tree fruit and nuts, which primarily involves shaking tree branches with an actuator at desired frequency and amplitude. Excitation of tree branches with vibratory signals causes the fruit to detach at the stem-fruit junction or branch-stem junction. In order to automate this process, the machine vision system needs to be capable of detecting branches in tree canopies and determine desirable shaking locations within detected branches to guide an actuator for harvesting. To achieve this goal, three specific research objectives were defined.

The first objective was to detect cherry tree branches using visible branch segments. Color-based image segmentation was used to detect branch sections that were visible to the canopy in the presence of full foliage during harvest season. Several morphological features of such branch sections including orientation, length, and width were used to link individual sections of the same branch together. The method achieved a branch detection accuracy of 89% in trees trained in vertical Upright Fruiting Offshoots (UFO) trellis architecture featuring branches ranging from
5 cm – 8 cm in diameter. This method was then tested in another cherry orchard trained in Y-trellis architecture with comparatively smaller branch diameter and denser foliage causing greater occlusion of tree branches. Only 55% branches were detected by the method in this condition. This led to the second objective to improve branch detection in dense foliage canopies.

The second objective was to use the positions of cherry clusters in the tree canopies as an additional information to detect branch location. Since a lot of branches or their sections are not visible to the camera in the presence of dense foliage and fruit during harvest season, it is difficult or sometimes even impossible to detect whole branches using only visible branch sections. The canopies may also include some branches that are too small to be detected. However, as cherries grow in and around branches, branch location could be estimated with reasonable accuracy using location and orientation of multiple cherry clusters. This method of detecting occluded branches using cherry position was successful especially in Y-trellis canopy architecture where majority of branches were less than 6 cm in diameter with presence of dense foliage and large clusters of cherries resulting in higher degree of branch occlusion. The integration of this method with the method described in previous paragraph resulted in on overall branch detection accuracy of 94% in Y-trellis architecture. For this architecture, use of cherry-cluster locations in the branch detection method improved detection accuracy by 39% compared to the method discussed in previous paragraph (Objective 1).

The third objective was aimed at developing a method for locating shaking positions in the tree branches for automated cherry harvesting. The localization of shaking position included the determination of shaking position on RGB images based on the cherry locations in different
canopy regions as well as estimation of the distance to the shaking position from the sensor. Depth information collected by 3D camera was mapped onto RGB image for distance estimation, with a root mean square error of 0.064 m. The performance of shaking method using automatically determined shaking positions was evaluated in Y-trellis and vertical canopy architectures. In Y-trellis architecture, a maximum fruit removal of 93% was achieved, which was 87% in vertical trellis architecture. It was also observed that two to three shaking positions per branch would remove most of the removable fruit in most of the branches in the two architectures evaluated in this study.

Overall, the results of this research showed huge potential of machine vision system to detect tree branches even in full foliage canopies during harvest season. It was also observed that branch detection could be improved substantially by utilizing cherry locations along with visible branch segments. A vision system using both color and 3D camera can be used to locate shaking positions in tree branches for shake-and-catch harvesting. Such machine vision system would be the key for developing fully automated cherry harvester with multiple shakers and catchers working at the same time. Multiple layered catching surfaces at different canopy levels have the potential to increase harvesting rate and decrease fruit damage.

6.2 Recommendations for Future Work

Through this research, several opportunities and challenges were identified for automating sweet cherry harvesting systems. In the following paragraphs, various areas for further research and development are discussed, which will help bring the automated cherry harvesting closer to the reality.
• This work focused on detecting whole tree branches by using clues from the segmented branch sections and cherry regions. The image segmentation results plays an important role in the accuracy of branch detection. Further improvement in image segmentation methods can play a role in improving the branch detection accuracy.

• The variability in lighting conditions can interfere with image processing algorithm leading to reduced branch detection accuracy. Therefore, it is recommended to keep consistent lighting condition throughout the imaging operation. A good lighting condition need to ensure proper illumination of the target object and at the same time avoid over illumination to reduce specular reflection form the surfaces.

• In this study, multiple cameras were used to capture the complete fruiting region of the tree. Using a camera with wider field of view can be beneficial for decreasing the complexity in imaging and image processing tasks for real-time application. A machine with multiple cameras, and multiple shaking arms and catching surfaces could also be investigated as a possible future direction for improving fruit quality along with harvesting efficiency. Multiple shakers and catching surfaces can also be developed to shake only smaller target regions to improve fruit removal, collection and quality measures through minimizing fruit drop height.

• Optimization of hardware and software for faster processing is also an important area of further investigation for field application of the technology.

• One of the challenges for detecting tree branches with desired level of accuracy is the tree architecture. A well trained tree canopy system is essential for developing robust algorithms for branch detection. Improvement in the uniformity on branch orientations
through proper training of branches to the trellis wires would help improve branch
detection and localization accuracy. Uniform and sufficient spacing between neighboring
branches could be another useful feature of a mechanization friendly architecture.
Appropriate tree training and pruning during both dormant and growing season may be
helpful in maintaining a thin fruiting wall architecture with short secondary branches,
which will potentially minimize the extent of branch occlusion.