BIOSYSTEMS ENGINEERING 146 (2016) 3–15

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/issn/15375110

Special Issue: Robotic Agriculture

Research Paper

Detection of cherry tree branches with full foliage in planar architecture for automated sweet-cherry harvesting

Suraj Amatya [a,b], *Manoj Karkee* [a,b,*], *Aleana Gongal* [a,b], *Qin Zhang* [a,b], *Matthew D. Whiting* [b,c]

[a] *Department of Biological Systems Engineering, United States*
[b] *Center for Precision and Automated Agricultural Systems, United States*
[c] *Department of Horticulture, Washington State University, United States*

ARTICLE INFO

Article history:
Received 27 March 2015
Received in revised form
2 October 2015
Accepted 14 October 2015
Published online 3 November 2015

Keywords:
Cherry harvesting
Image segmentation
Bayesian classifier
Branch detection
Vertical planar architecture
Upright fruiting offshoots

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.

* *Corresponding author*. 24106 N Bunn Rd, Prosser, WA 99350, United States. Tel.: +1 509 786 9208.
E-mail address: manoj.karkee@wsu.edu (M. Karkee).

http://dx.doi.org/10.1016/j.biosystemseng.2015.10.003
1537-5110/

Nomenclature

USDA	United States Department of Agriculture
UFO	Upright Fruiting Offshoots
RDA	Rapid Displacement Actuator
3D	3-Dimensional
2D	2-Dimensional
WSU	Washington State University
mm	Millimetres
LED	Light Emitting Diode
HFOV	Horizontal Field of View
RGB	Red–Green–Blue
CIELAB	1976 CIE L*a*b*
x	Feature vector
wi	Class definition
P(wi)	Prior probability of class wi
pdf	Probability density function
p(x|wi)	Class conditional distribution of x given that it belongs to class wi
p(wi|x)	Posterior probability of wi given the evidence x
p(x)	Probability of x
di(x)	Decision function to decide class given the value of x
Ci	Covariance matrix
mi	Mean vector
|Ci|	Determinant of covariance matrix
ln	Natural logarithm
px	Pixels
Y	Pixel position in rows
X	Pixel position in columns
m	Slope
C	Intercept

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, Freeborn, & Long, 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 (Fennimore & Doohan, 2008; Gongal, Amatya, Karkee, Zhang, & Lewis, 2015; Hertz & 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 (Long, 2010; Peterson & Wolford, 2001) 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 (Chen et al., 2012; Du, Chen, Zhang, Scharf, & Whiting, 2011). An economic study suggested that mechanically harvested cherry production systems will return more money to growers than the traditional system (Seavert & Whiting, 2011). The study was based on the results from a USDA mechanical harvester evaluated in early 2000's (Peterson & 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, Whiting, & Wolford, 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 & 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 (Du, Chen, Zhang, Scharf, & Whiting, 2012; Halderson, 1966; Larbi & Karkee, 2014; Norton et al., 1962; Peterson et al., 2003; Peterson & Wolford, 2001; Zhou, He, Zhang, & Karkee, 2014), 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, Fujiura, Akase, & Imagawa, 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, Razdan, Femiani, Cui, & Wonka, 2007; Laptev et al., 2000; Trinder & Wang,

1998). Morphological features has been widely used in these studies for detecting and joining road segments to extract the road network (Valero, Chanussot, Benediktsson, Talbot, & Waske, 2010; Zhang, Murai, & Baltsavias, 1999). Lü, Cai, Liu, Deng, and Zhang (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, Tisseyre, Sinfort, Tillet, and Sevila (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, Du, Qiu, and Wu (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 & Lu, 2006; Karkee, Adhikari, Amatya, & Zhang, 2014; Karkee & 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.

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 subsections.

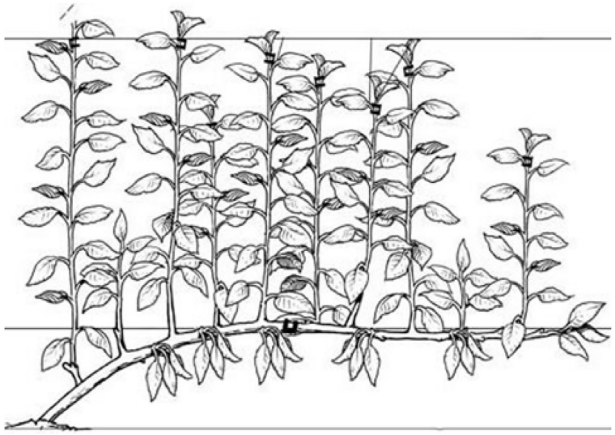

Fig. 1 – Simplified representation of a cherry tree canopy trained in upright fruiting offshoots (UFO) System.

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 (Fig. 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 (Fig. 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

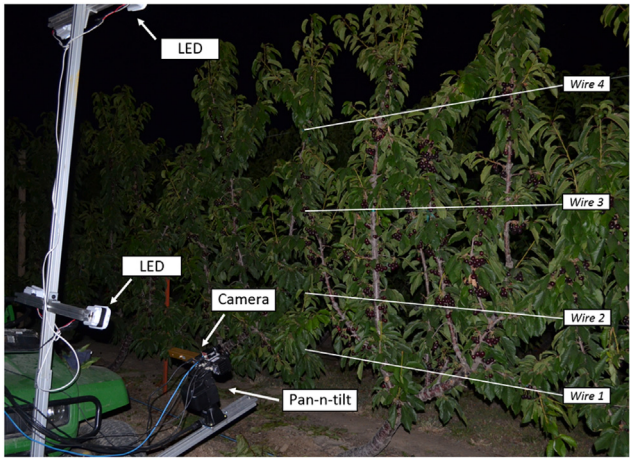

Fig. 2 – Field image acquisition setup with artificial lighting. Camera mounted on a pan-n-tilt system took three images capturing region between wire 1 and wire 4.

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 stereo-vision 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 (Fig. 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.

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 (Fig. 3).

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, Kar, & Bhattacharya, 2011). Figure 4 shows an example of removal of specular reflection from the image.

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 form 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.

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 & Stockman, 2001, Eq (1)). The input parameters for Bayesian classifier is a multivariate feature vector x for all classes w_i. The prior probability,

Fig. 3 – a) Original image; and b) image after image enhancement by increasing contrast ratio.

Fig. 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.

$P(w_i)$, 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, Hart, & Stork, 1973) as follows.

$$p(w_i|x) = \frac{p(x|w_i)*P(w_i)}{p(x)} \quad (1)$$

Where,

$$p(x) = \sum_{i=1}^{n} p(x|w_i)P(w_i) \quad (2)$$

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:

$$d_i(x) = p(x|w_i)*P(w_i) \quad (3)$$

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

$$p(x|w_i) = \frac{1}{(2\pi)^{\frac{n}{2}}|C_i|^{\frac{1}{2}}} exp^{-\frac{1}{2}\left[(x-m_i)^T C_i^{-1}(x-m_i)\right]} \quad (4)$$

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) takes the following form (Gonzalez, Woods, & Eddins, 2009).

$$d_i(x) = ln\, P(w_i) - \frac{1}{2}ln|C_i| - \frac{1}{2}\left[(x-m_i)^T C_i^{-1}(x-m_i)\right] \quad (5)$$

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 Fig. 5. For ellipses, two axes are defined; the major axis and the minor axis which is the axis perpendicular to the major (Fig. 5).

Steps involved in branch detection are shown in the flowchart in Fig. 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.

Most of the branches of the cherry trees used this study were approximately 50–60 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

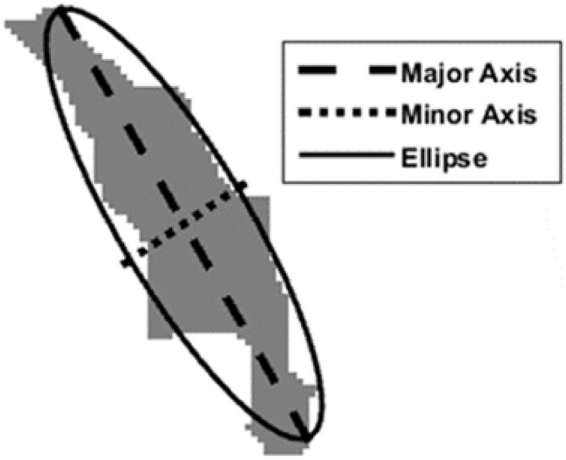

Fig. 5 – Equivalent ellipse, major axis length and minor axis length of a segmented region (shaded region).

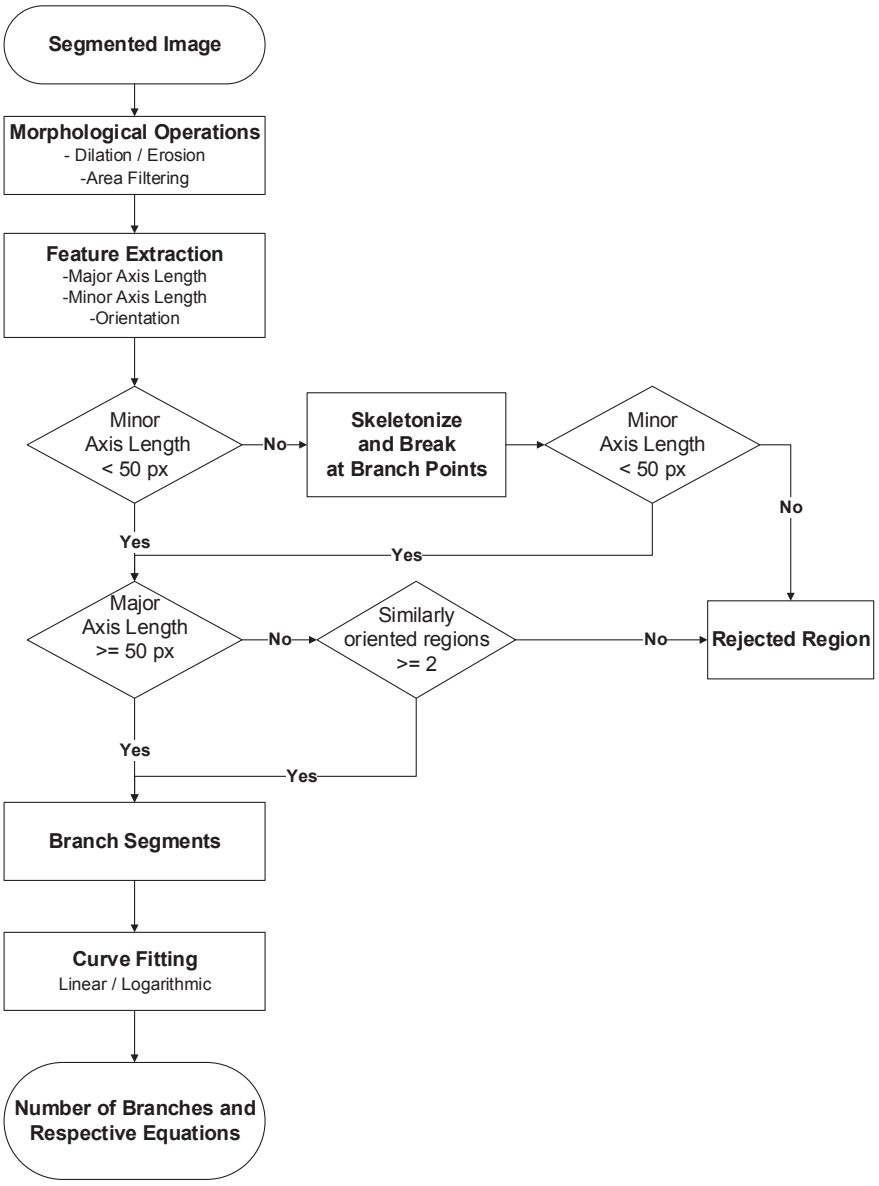

Fig. 6 – Flowchart of the branch detection method. The input to the process is a segmented cherry tree image and the output is the number of branches in an image and equations representing those branches.

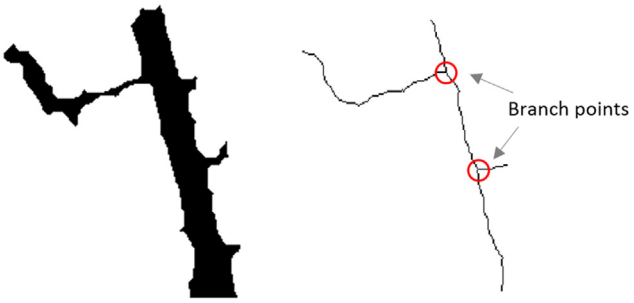

Fig. 7 – Skeletonisation and fragmentation of cherry branch segments. Medial axis skeleton was obtained and branching points were identified in the skeleton. Branches were then divided at branching points to form smaller, individual sections.

until only the skeleton remains (Lam, Lee, & Suen, 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 (Fig. 7). Individual sections were later connected to other appropriate sections to form longer branch segments.

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 image captured and it can be assumed that the branches extend well above and below the current field of view. In Fig. 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 20° of the orientation of the segment of interest, the segments were identified as parts of the same branch. All four sections shown in Fig. 8 had similar orientation (i.e. relative orientation within 20°), 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.

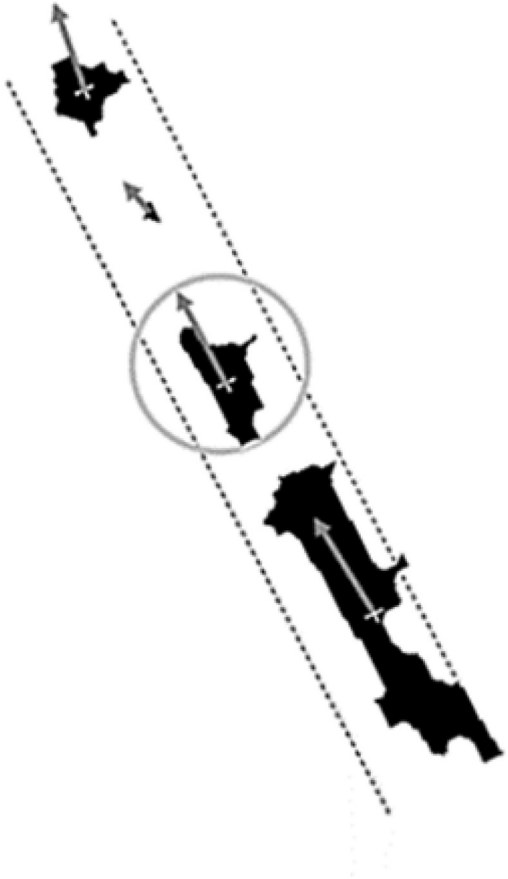

Fig. 8 – Orientations of branch segments represented by arrows. Two regions were assumed to be parts of the same branch when their relative orientations were within 20°.

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,

$$\text{Linear Model}: \quad Y = m^{*}X + C \tag{6}$$

$$\text{Logarithmic Model}: \quad Y = m^{*}\log(X) + C \tag{7}$$

where, Y = pixel position in rows

X = pixel position in columns
m = slope of equation
C = 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. Results and discussions

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).

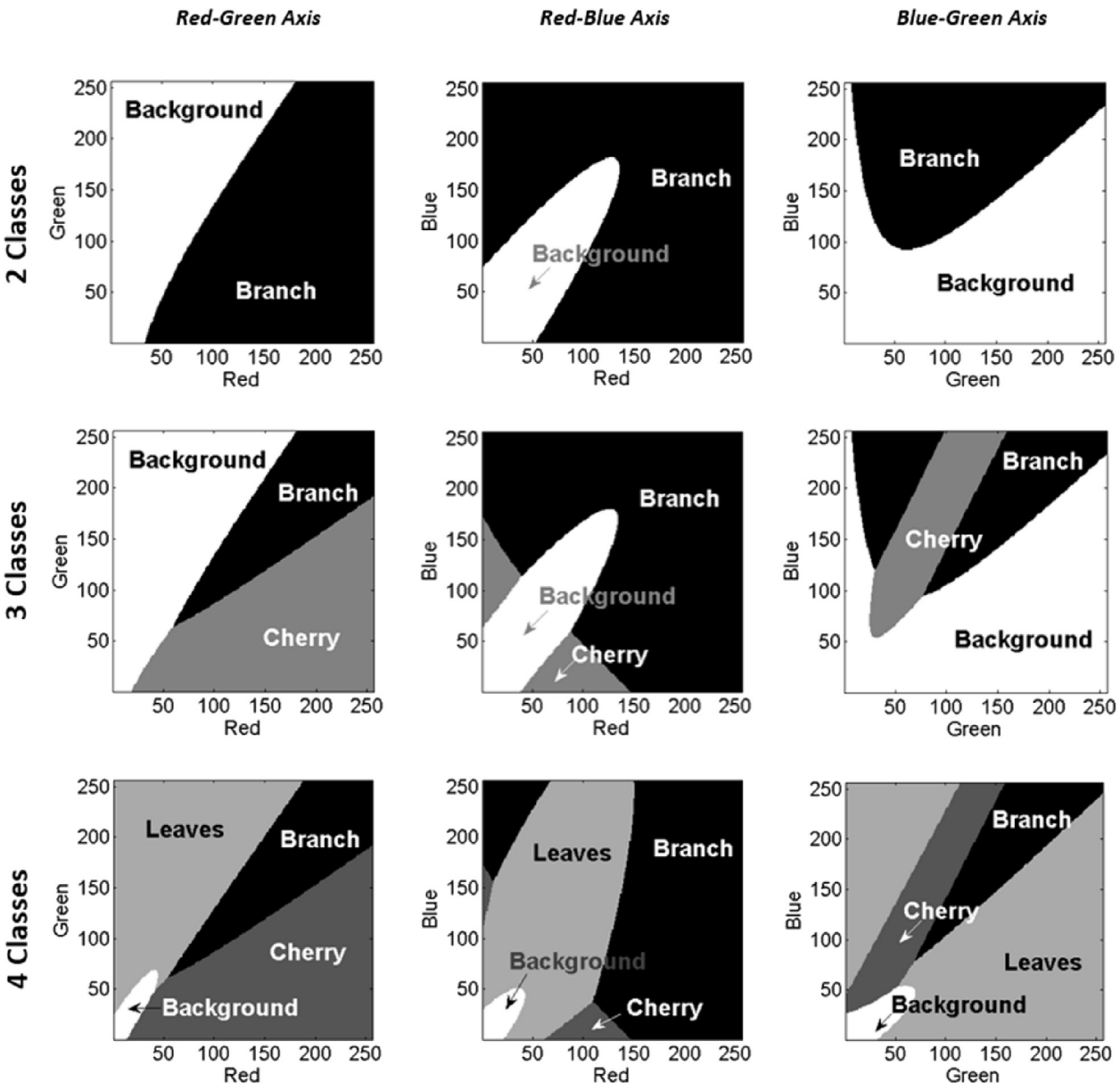

Fig. 9 – Decision boundaries for different classes of pixels in R–G, R–B, and B–G planes of a RGB colour space.

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 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 Fig. 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 (Fig. 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.

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 1) and 72.4% (Table 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 1) for two class classification was due to the misclassification of cherry pixels into branch class as evident from the decision

Table 1 – Classification accuracy for two-, three-, and four-class classification with training dataset.

Class	Training set								
	2 classes			3 classes			4 classes		
	Accuracy (%)	False positives (%)	False negatives (%)	Accuracy (%)	False positives (%)	False negatives (%)	Accuracy (%)	False positives (%)	False negatives (%)
Branch	75.6	24.4	13.4	87.0	13.0	15.0	91.0	9.0	13.2
Cherry	–	–	–	87.2	12.8	50.5	73.0	27.0	38.7
Leaf	–	–	–	–	–	–	90.9	9.1	6.8
Background	98.7	1.3	2.8	96.0	4.0	1.3	91.2	8.8	8.2

Table 2 – Classification accuracy for two-, three-, and four-class classification with testing dataset.

Class	Testing set								
	2 classes			3 classes			4 classes		
	Accuracy (%)	False positives (%)	False negatives (%)	Accuracy (%)	False positives (%)	False negatives (%)	Accuracy (%)	False positives (%)	False negatives (%)
Branch	72.4	27.6	11.6	84.3	15.7	14.1	89.6	10.4	12.4
Cherry	–	–	–	79.7	20.3	53.9	73.3	26.7	43.3
Leaf	–	–	–	–	–	–	86.9	13.1	5.4
Background	98.8	1.2	3.5	95.6	4.4	1.9	91.2	8.7	12.6

boundaries (Fig. 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 (Fig. 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%–90% for testing dataset). Table 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) is the fraction of correctly classified pixels with respect to all the pixels belonging to that class.

The sample result of Bayesian segmentation is shown in Fig. 10 where image pixels are classified as either branch, cherry, leaf, or background region based on their colour features.

Table 3 – Confusion matrix of testing set for four class classification.

		Actual class (pixels)				Consumer accuracy
		Branch	Cherry	Leaf	Background	
Predicted class (pixels)	Branch	302,731 (89.6%)	8480 (2.5%)	21,059 (6.2%)	5761 (1.7%)	89.6%
	Cherry	16,155 (10.1%)	117,252 (73.3%)	1940 (1.2%)	24,679 (15.4%)	73.3%
	Leaf	18,546 (1.2%)	1221 (0.1%)	1,360,769 (86.9%)	185,132 (11.8%)	86.9%
	Background	8122 (0.5%)	79,690 (4.9%)	55,041 (3.4%)	1,492,162 (91.2%)	91.2%
	Producer accuracy	87.6%	56.7%	94.6%	87.4%	

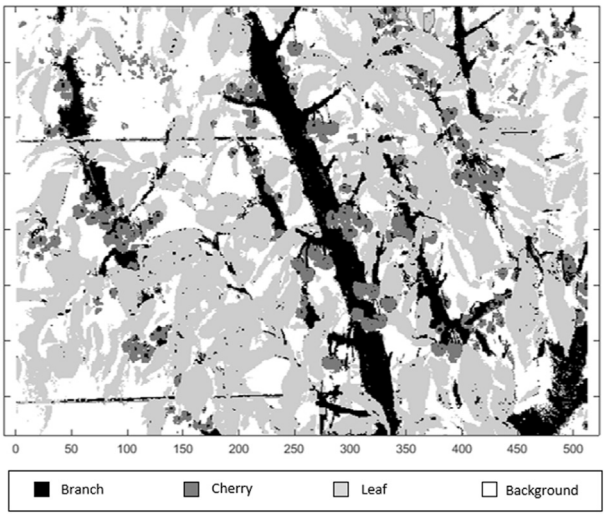

Fig. 10 – Sample result of Bayesian classification of an image to branch, cherry, leaf and background classes.

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 braches 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 braches 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 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 pre-processing 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.

Figure 11 shows example results from the branch detection algorithm in which each detected branch has been plotted over the original image (Fig. 11b and d). Corresponding original images are shown in Fig. 11a and c. 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 (He et al., 2012b; Zhou et al., 2014, 2012). 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 braches 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.

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 (de Bergh & Gool, 2011; Alenyà, Dellen, & Torras, 2011; Gongal, Silwal, et al., 2015).

Table 4 – Accuracy in detecting individual cherry branches.

	Camera position			
	Low	Medium	High	Overall
Detected (%) (no. of branches)	87.4% (139)	92.1% (139)	88.1% (126)	89.2% (404)
False positives (%) (no. of branches)	18.9% (30)	15.2% (23)	14.0% (20)	16.1% (73)
Undetected (%) (no. of branches)	12.6% (20)	7.9% (12)	11.9% (17)	10.8% (49)

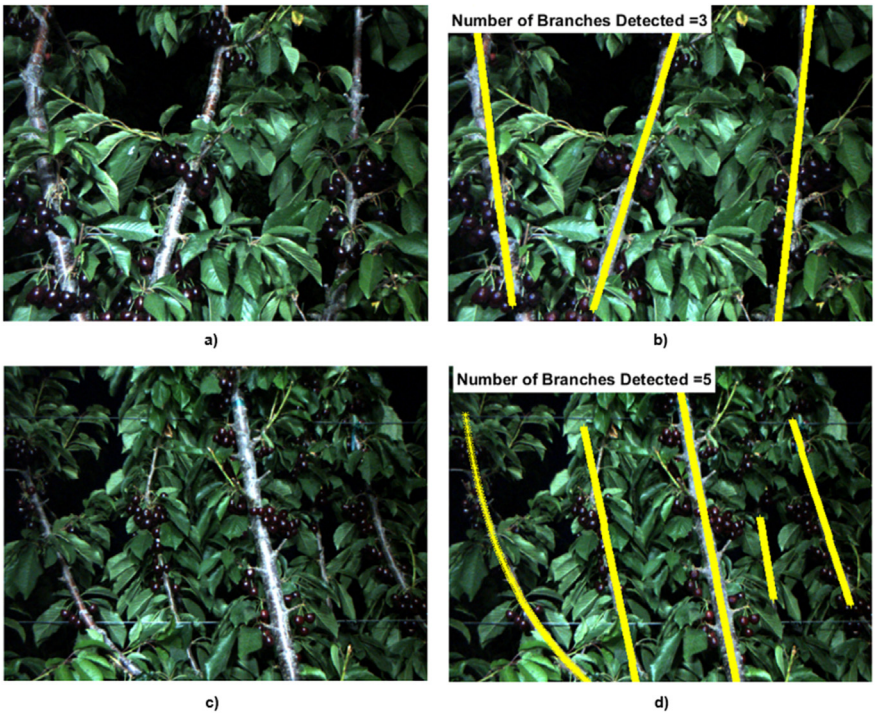

Fig. 11 – Individual branch detection: original image (left); detected branches (right).

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 & 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, Silwal, et al., 2015, 2014; Silwal, Gongal, & Karkee, 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.

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

Alenyà, G., Dellen, B., & Torras, C. (2011). 3D modelling of leaves from color and ToF data for robotized plant measuring. In 2011

IEEE international conference on robotics and automation. Shanghai, China, May 9–13.
de Bergh, M. V., & Gool, L. V. (2011). *Combining RGB and ToF cameras for real-time 3D hand gesture interaction* (pp. 66–72). Applications of Computer Vision (WACV).
Chen, D., Du, X., Zhang, Q., Whiting, M., Scharf, P., & Wang, S. (2012). Performance evaluation of mechanical cherry harvesters for fresh market grade fruits. *Applied Engineering in Agriculture, 28*(4), 483–489.
Das, A., Kar, A., & Bhattacharya, D. (2011). Elimination of specular reflection and identification of ROI: the first step in automated detection of cervical cancer using digital colposcopy. In *IEEE international conference on imaging systems and techniques (IST)* (pp. 237–241).
Du, X., Chen, D., Zhang, Q., Scharf, P. A., & Whiting, M. D. (2011). *Mechanical harvesting of UFO cherry: Investigation of tree plant dynamics*. ASABE Paper No. 1110523. St. Joseph, Mich.: ASABE.
Du, X., Chen, D., Zhang, Q., Scharf, P. A., & Whiting, M. D. (2012). Dynamic responses of sweet cherry trees under vibratory excitations. *Biosystems Engineering*, 305–314 (2012).
Duda, R. O., Hart, P. E., & Stork, D. G. (1973). *Pattern classification* (2nd ed.). A Wiley-Interscience Publication. ISBN 0-471-05669-3.
Employment Security Dept. (2013). *2012 agricultural workforce report. Labor and market performance analysis*. Washington State: Employment Security Department.
Fennimore, S. A., & Doohan, D. J. (2008). The challenges of specialty crop weed control, future directions. *Weed Technology, 22*, 364–372.
Gao, M., & Lu, T. F. (2006). Image processing and analysis for autonomous grapevine pruning. Mechatronics and automation. In *Proceedings of the 2006 IEEE international conference, 922 – 927*.
Gongal, A., Amatya, S., & Karkee, M. (2014). *Identification of repetitive apples for improved crop-load estimation with dual-side imaging*. ASABE, Paper Number: 141888882, St. Joseph, MI.
Gongal, A., Amatya, S., Karkee, M., Zhang, Q., & Lewis, K. (2015). Sensors and systems for fruit detection and localization: a review. *Computer and Electronics in Agriculture, 116*(2015), 8–19.
Gongal, A., Silwal, A., Amatya, S., Karkee, M., Zhang, Q., & Lewis, K. (2015). Apple crop-load estimation with over-the-row machine vision system. *Computers and Electronics in Agriculture*. Accepted for Publication.
Gonzalez, R. C., Woods, R. E., & Eddins, S. L. (2009). *Digital image processing using MATLAB* (2nd ed.). Upper Saddle River, NJ: Prentice Hall.
Halderson, J. L. (1966). *Fundamental factors in mechanical cherry harvesting*. Transactions of the ASAE. ASAE Paper No. 63– 120.
He, L., Du, X., Qiu, G., & Wu, C. (2012a). *3D reconstruction of Chinese hickory trees for mechanical harvest*. ASABE Paper No. 12-1340678. St. Joseph, Mich.: ASABE.
Hertz, T., & Zahniser, S. (2013). Is there a farm labor shortage? *American Journal of Agricultural Economics, 95*(2), 476–481.
He, L., Zhou, J., Du, X., Chen, D., Zhang, Q., & Karkee, M. (2012b). *Shaking energy delivery on sweet cherry trees in different excitation models*. ASABE Paper No. 12-1337766. St. Joseph, Mich.: ASABE.
Hu, J., Razdan, A., Femiani, J. C., Cui, M., & Wonka, P. (2007). Road network extraction and intersection detection from aerial images by tracking road footprints. *IEEE Transactions on Geoscience and Remote Sensing, 45*(12), 4144–4157.
Karkee, M., & Adhikari, B. (2015). A method for three dimensional reconstruction of apple trees for automated pruning. *Transactions of the ASABE, 58*(3), 565–574.
Karkee, M., Adhikari, B., Amatya, S., & Zhang, Q. (2014). Identification of pruning branches in tall spindle apple trees for automated pruning. *Computer and Electronics in Agriculture, 103*(2014), 127–135.
Lam, L., Lee, S. W., & Suen, C. Y. (1992). Thinning methodologies – a comprehensive survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 14*(9), 869–885.
Laptev, I., Mayer, H., Lindeberg, T., Eckstein, W., Steger, C., & Baumgartner, A. (2000). Automatic extraction of roads from aerial images based on scale space and snakes. *Machine Vision and Applications, 12*(2000), 23–31.
Larbi, P. A., & Karkee, M. (2014). Effects of orchard characteristics and operator performance on harvesting rate of a mechanical sweet cherry harvester. *GSTF Journal on Agricultural Engineering, 1*(1), 1–11.
Long, L. E. (2010). *Worldwide trends in cherry training systems*. Available at: http://extension.oregonstate.edu (Accessed on 04.14.15).
Lü, Q., Cai, J., Liu, B., Deng, L., & Zhang, Y. (2014). Identification of fruit and branch in natural scenes for citrus harvesting robot using machine vision and support vector machine. *International Journal of Agriculture and Biological Engineering, 7*(2), 115–121.
McFarlane, N. J. B., Tisseyre, B., Sinfort, C., Tillet, R. D., & Sevila, F. (1997). Image analysis for pruning of long wood grape vines. *Journal of Agricultural Engineering Research, 66*(1997), 111–119.
Norton, R. A., Claypool, L. L., Leonard, S. J., Adrian, P. A., Fridley, R. B., & Charles, F. M. (1962). Mechanical harvesting of sweet cherries. *California Agriculture, 16*(5), 8–10.
Peterson, D. L., Whiting, M. D., & Wolford, S. D. (2003). Fresh-market quality tree fruit harvester part I: sweet cherry. *Applied Engineering in Agriculture, 19*(5), 539–543.
Peterson, D. L., & Wolford, S. D. (2001). Mechanical harvester for fresh market quality stemless sweet cherries. *Transactions of the ASAE, 44*(3), 481–485.
Seavert, C., Freeborn, J., & Long, L. (2008). *Orchard economics: Establishing and producing high-density sweet cherries in Wasco County*. OSU Extension Service Publication. EM 8802-E.
Seavert, C. F., & Whiting, M. D. (2011). *Comparing the economics of mechanical and traditional sweet cherry*. Acta Hort. (ISHS) 903:725-730.
Shapiro, L. G., & Stockman, G. C. (2001). *Computer vision*. Upper Saddle River, New Jersey: Prentice-Hall, Inc. ISBN 0-13-030796-3.
Silwal, A., Gongal, A., & Karkee, M. (2014). Apple identification in field environment with over-the-row machine vision system. *Agricultural Engineering International: Agric Eng Intl (CIGR Journal), 16*(4), 66–75.
Tanigaki, K., Fujiura, T., Akase, A., & Imagawa, J. (2008). Cherry harvesting robot. *Computer and Electronics in Agriculture, 63*(2008), 65–72.
Trinder, J. C., & Wang, Y. (1998). Automatic road extraction from aerial images. *Digital Signal Processing, 8*(1998), 215–224.
USDA. (2013). *2013 Washington annual agriculture bulletin*. Washington Field Office: USDA/National Agriculture Statistics Service.
Valero, S., Chanussot, J., Benediktsson, J. A., Talbot, H., & Waske, B. (2010). Advanced directional mathematical morphology for the detection of the road network in very high resolution remote sensing images. *Pattern Recognition Letters, 31*(2010), 1120–1127.
Whiting, M. D. (2009). *Upright fruiting offshoots*. WA: WSU-IAREC, Prosser. Available at: http://cahnrs-cms.wsu.edu/StoneFruit/research/Pages/ufo.aspx (Accessed on 04.14.15).
Young, C., & Kupferman, E. (2015). *In-field hydrocooling cherry temperature management*. WSU-TFREC, Postharvest Information Network. Available at: http://postharvest.tfrec.wsu.edu/pages/J5I1D (Accessed on 04.14.15).
Zhang, C., Murai, S., & Baltsavias, E. (1999). Road network detection by mathematical morphology. In *Proc. ISPRS workshop on 3D geospatial data production: Meeting application requirements, Paris* (pp. 185–200).

Zhou, J., He, L., Du, X., Chen, D., Zhang, Q., & Karkee, M. (2012). *Dynamic response of sweet cherry tree to the vibration of a limb shaker*. ASABE Paper No. 12-1337429. St. Joseph, Mich.: ASABE.

Zhou, J., He, L., Zhang, Q., & Karkee, M. (2014). Effect of excitation position of a handheld shaker on fruit removal efficiency and damage in mechanical harvesting of sweet cherry. *Biosystems Engineering*, *125*, 36–44.