Computers and Electronics in Agriculture 120 (2016) 26–35

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Original papers

Apple crop-load estimation with over-the-row machine vision system

A. Gongal [a,b], A. Silwal [a,b], S. Amatya [a,b], M. Karkee [a,b,*], Q. Zhang [b], K. Lewis [b,c]

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

ARTICLE INFO

Article history:
Received 11 August 2014
Received in revised form 20 October 2015
Accepted 27 October 2015
Available online 3 December 2015

Keywords:
Crop-load estimation
Machine vision
Apple identification
Occlusion
3D registration

ABSTRACT

Accurate crop-load estimation is important for efficient management of pre- and post-harvest operations. This information is crucial for the planning of labor and equipment requirement for harvesting and transporting fruit from the orchard to packing house. Current machine vision-based techniques for crop-load estimation have achieved only limited success mostly due to: (i) occlusion of apples by branches, leaves and/or other apples, and (ii) variable outdoor lighting conditions. In order to minimize the effect of these factors, a new sensor system was developed with an over-the-row platform integrated with a tunnel structure which acquired images from opposite sides of apple trees. The tunnel structure minimized illumination of apples with direct sunlight and reduced the variability in lighting condition. Images captured in a tall spindle orchard were processed for identifying apples, which achieved an identification accuracy of 79.8%. The location of apples in three-dimensional (3D) space was used to eliminate duplicate counting of apples that were visible to cameras from both sides of the tree canopy. The error on identifying duplicate apples was found to be 21.1%. Overall, the method achieved an accuracy of 82% on estimating crop-load on trees with dual side imaging compared to 58% with single side imaging. Over-the-row machine vision system showed promise for accurate and reliable apple crop-load estimation in the apple orchards.

1. Introduction

Crop-load estimation is essential for efficient and effective management of orchards at various stages during the production cycle. In early summer, crop-load estimation (including counting and location of apples) helps to optimize green fruit thinning (Volz, 1988), which is critically important to improve fruit quality and yield. With this information, a producer can develop more sound crop load management and harvest strategies and can maximize their profits (Mizushima and Lu, 2011). It also helps producers to insure their crops (Cohen et al., 2011) so that they can be compensated in the event of financial loss due to weather related crop loss. Further, crop-load estimation prior to harvest is essential for efficient management of the labor force, harvest equipment, and vehicles for transportation of fruit from field to the packing plant. Packing house can also be benefited by optimizing postharvest handling process and storage with early crop-load information (Cohen et al., 2011).

Accurate crop-load estimation has always been a challenge for tree fruit producers (Winter, 1986). Several forecasting models have been developed based on ecological and/or cultivar related parameters for yield prediction (Stajnko et al., 2004). One such example is the 'Prognosfruit' crop forecasting model, which estimates the number of apples per tree based on yield capacity of tree, fruit density and average fruit mass (Winter, 1986). This method is time consuming, as different parameters have to be measured or estimated for individual orchards (Stajnko et al., 2004). Researchers have also investigated the development of yield forecasting models based on vegetation indices estimated with hyperspectral or multispectral images captures by areal platforms (Best et al., 2008; Ye et al.,2007). Ye et al. (2007) predict citrus yield from hyperspectral images using vegetatition indices (e.g. NDVI, SR and PRI) and Partial Least Square (PLS) regression model. However the accuracy of these methods has been limited due to climate variabilities, cultivars and geographic locations (Aravena Zamora et al., 2010). Crop-load estimation based on direct counting of fruits in selected trees is another approach. However, manual countings are generally based on random sampling within trees, which require high labor input. Manual sampling may also have low statistical efficiency and could be biased (Aravena Zamora et al., 2010). In order to avoid such a bias on estimation, Wulfsohn et al. (2012) classified the tree rows based on vegetative indices to develop multi-level systematic (manual) sampling techniques to estimate fruit counts. But, manual sampling is time

* Corresponding author at: 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.compag.2015.10.022
0168-1699/

consuming and has chances of inaccuracy due to manual error (Linker et al., 2012).

To address these issues, machine vision systems involving different types of sensors and image processing techniques have been investigated. Aggelopoulou et al. (2011) used a color camera to identify flowers in trees, which was then correlated to crop-load. In such a system, fruit set, thinning and fruit drop later in the season (Winter, 1986) will complicate the estimation and affect the accuracy of estimation. To estimate crop-load based on identification of individual fruit, various studies have been carried out using different types of sensors like color camera (Linker et al., 2012; Tabb et al., 2006), thermal camera (Stajnko et al., 2004), and multispectral and hyperspectral cameras (Kim and Reid, 2004; Best et al., 2008; Safren et al., 2007). Further, various types of image processing techniques have been investigated to identify fruit, which can roughly be divided into shape-based and spectral-based analysis (Bulanon and Katoka, 2010). Bulanon et al. (2002) used spectral analysis to identify apples and reported 88% accuracy. Linker et al. (2012), Stajnko et al. (2009), and Ji et al. (2012) used shape, color and texture based analysis and reported 85%, 89%, and 89% accuracy respectively. Fruit detection accuracy in those studies was limited by clustering and occlusions of fruits (Stajnko et al., 2009), and variable lighting conditions in orchards (Ji et al., 2012; Linker et al., 2012).

Most of these studies, however, evaluated the accuracy based on the number of apples visible in an image taken from one side of the tree canopy, which may be substantially different from the total number of apples in a tree. To address this issue, Wang et al. (2013) used a stereovision camera at night for multiple side imaging and compared the results with the actual number of apples counted by humans. This study achieved 60% accuracy in crop-load estimation when substantial number of apples were in clusters. Similarly, Vision Robotics Corp (San Diego, CA) reported an accuracy of 40–60% without correction for non-visible apples in images (Koselka, 2010). To improve the accuracy, the study correlated the manual apple count with the vision-based estimated count. Zhou et al. (2012) also identified red and green apple using color information and correlated that to the ground truth with a correlation coefficient of 0.58–0.71. However, this process will be labor intensive and specific to particular type of orchard. Correlation coefficient may also change from year to year. Therefore, it is valuable to develop a machine vision system that can improve crop-load estimation accuracy to an acceptable level without correlating estimated count and ground truth collected by human.

Fruit identification in an orchard environment through image analysis is affected by various factors including occlusion of fruit, variable outdoor lighting conditions, complex plant structures, and irregularity of fruit shape and size (Cohen et al., 2011; Karkee and Zhang, 2012). As discussed before, the two most critical factors limiting the accuracy of crop-load estimation in apple orchards are; (i) a significant number of apples are invisible to cameras looking from only one side of apple tree canopies; and (ii) apple identification accuracy is compromised by variable lighting conditions. It was observed in the field that only 60% of apples were visible from one side of the canopy even in a modern tall spindle apple orchard because of occlusion of apples by leaves, stems and other apples. In addition, apples can be partially occluded or can be in clusters of two or more fruit, both situations are challenging for image processing techniques. Varying lighting conditions, on the other hand, causes non-uniform distribution of light intensity in apples based on their exposure to sunlight at the time of imaging, which may cause improper and incomplete image segmentation. Thus, it would be challenging to develop a robust algorithm to identify apples in variable lighting conditions. In this study, our focus is to minimize the effect of variable lighting conditions and occlusion of apples to improve apple crop-load estimation in orchard environment. The major goal of this study was to investigate the potential of improving crop-load estimation with dual-side imaging compared to single-side imaging. The specific objectives of the study were:

1. To identify and count total number of apples in the images captured from opposite sides of tree canopies using an Over-the-Row (OTR) sensor platform; and
2. To use 3D location information to estimate number of apples visible from both sides and to avoid duplicate counting.

2. Materials and methods

An over-the-row (OTR) sensing platform with dimensions 2.13 m (length) × 2.74 m (width) × 3.67 m (height) was developed and mounted on the three-point hitch of a tractor so that it could be lifted up from one trees to another and dropped down to the ground while taking images. A color camera (Fig. 1a) (Prosilica GigE 1290c, Allied Vision Technologies, Stadtroda, Germany) with a field of view (FOV) of 43.6° (horizontal) by 33.4° (vertical) and image resolution of 1280 × 960 was used to acquire RGB (color) images. Also, a time-of-flight-of-light-based 3D camera (Fig. 1a) (PMD CamCube 3.0, PMD Technologies, Siegen, Germany) with image resolution of 200 × 200 and FOV of 40° (horizontal) and 40° (vertical) was used to obtain 3D information of objects. Two cameras were integrated together and mounted in the middle (lengthwise) of the platform with sliding mechanism so that camera height could be adjusted to capture the whole tree. To block direct sunlight while imaging, a tunnel (Fig. 1b) was created by covering the platform with tarpaulin (opaque curtain) on four sides and the top of the platform. A number of LED lights (Fig. 1c) (Trilliant® 36 Light Emitting Doide Grote, Madison, Indiana) were installed in the platform to create a controlled, uniform lighting environment inside the tunnel. The lighting system also added capability for nighttime data collection.

Four major steps were involved in identifying and counting apples while avoiding duplicate counting: (i) data collection, (ii) apple identification, (iii) co-registration of two dimensional (2D) color images and 3D images, and (iv) identification of duplicate apples and apple counting (Fig. 2). These steps are described in details in the following subsections.

2.1. Data collection

Canopy images were acquired from two sides of a row in commercial apple orchards using an over-the-row sensing platform. Apple trees used in this work were of 'Jazz' variety (Fig. 3) trained in the tall spindle architecture (row spacing 2.74 and inter-plant spacing 1.17 m) (Yakima Valley Orchards, Prosser, WA). Images were acquired between four and one week before harvest season in 2013. During the earlier weeks of data collection, the apples were immature and green in color. During the last week of imaging, apples matured and acquired their full size and red color with good hue and intensity. The color and 3D cameras used in this work had relatively small field-of-view, which were not enough to capture the whole tree canopy at once. To cover entire canopies, images were acquired from five different camera heights (Fig. 4). The over-the-row platform with sensing system was driven through apple tree rows to capture images of 20 trees. A total of 212 canopy images were captured with 3D camera and another 212 images were captured by the color camera (424 images in total). Out of 212 color images, 104 images were captured when apples were still imagure and 108 images were captured from trees with mature fruit. Two people counted the total number of apples in each tree manually for the ground truth dataset. Actual number of apples that were visible in the images captured from both sides