Computers and Electronics in Agriculture 116 (2015) 8–19

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

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

Review

Sensors and systems for fruit detection and localization: A review

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

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

ARTICLE INFO

Article history:
Received 31 October 2014
Received in revised form 30 May 2015
Accepted 30 May 2015
Available online 18 June 2015

Keywords:
Machine vision
Fruit detection
Fruit localization
Robotic harvesting
Crop-load estimation
Specialty crops

ABSTRACT

This paper reviews the research and development of machine vision systems for fruit detection and localization for robotic harvesting and/or crop-load estimation of specialty tree crops including apples, pears, and citrus. Variable lighting condition, occlusions, and clustering are some of the important issues needed to be addressed for accurate detection and localization of fruit in orchard environment. To address these issues, various techniques have been investigated using different types of sensors and their combinations as well as with different image processing techniques. This paper summarizes various techniques and their advantages and disadvantages in detecting fruit in plant or tree canopies. The paper also summarizes the sensors and systems developed and used by researchers to localize fruit as well as the potential and limitations of those systems. Finally, major challenges for the successful application of machine vision system for robotic fruit harvesting and crop-load estimation, and potential future directions for research and development are discussed.

Contents

* 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.2015.05.021

1. Introduction

Harvesting of specialty crops such as apples, citrus, cherries and pears is highly labor intensive and is becoming less sustainable with increasing cost and decreasing availability of a skilled labor force. In Washington State alone, more than 15 billion apples have to be handpicked by seasonal labor with an estimated harvesting cost of \$1150–\$1700 per acre per year (Gallardo et al., 2010). Further, hand harvest activities pose high risk of back strain and musculoskeletal problems to fruit pickers due to repetitive hand motions, awkward postures while picking fruit at high locations, and ascending, and descending on ladders with heavy loads (Fathallah, 2010). Hofmann et al. (2006) found that \$21 million compensation was paid for ladder-related injuries in the Washington State tree fruit industry between 1996 and 2001, which accounted for almost half of all compensable claims in that sector for the six-year period. Projecting into the future, the labor issue is expected to become more critical in terms of both increasing costs and uncertain availability (Fennimore and Doohan, 2008). Thus, automated or robotic harvesting system is essential to meet the increasing labor demand, to lower human risk of injuries in orchards, and to decrease the harvesting cost by saving time, money, and energy, which is profitable to both producers and consumers.

Schertz and Brown (1968) first proposed the concept of automated harvesting as an alternate to mechanical harvesting. Sistler (1987) argued that research and development on machine vision, robotics, and intelligent machine is key to improve the sustainability of specialty crop production in USA. Pejsa and Orrock (1984) and Sistler (1987) indicated that citrus fruit and apple has greatest potential for fruit detection and robotic harvesting. However, there have been several challenges in the successful development of such technologies; (i) lack of a clear direction for agricultural automation and robotics among researchers, commodity groups, and governments; (ii) variable and uncertain outdoor environment; (iii) complex plant structure; (iv) variable product shape and size; and (v) lack of support system for repair and maintenance of equipment (Sistler, 1987). Although no robotic or automated system has yet been commercialized for harvesting fresh-market tree fruit crops, various researchers, and private companies, have attempted fruit detection, localization, and robotic harvesting from different perspectives. Li et al. (2011) recognized that substantial improvement in fruit detection and localization accuracies is necessary to practice robotic harvesting. Detection and localization of fruit are the most fundamental information for machine vision based harvester. Efforts in improving fruit detection and localization will also help to improve accuracy of crop-load estimation.

Numerous research efforts have been reported in the literature on the development of a machine vision system for image acquisition, fruit detection, and fruit localization for robotic harvesting of fruits (Parish and Goksel, 1977; Harrell et al., 1985; Yang et al., 2007; Baeten et al., 2008; Scarfe et al., 2009; Bulanon et al., 2010; Yamamoto et al., 2014). However, proper synthesization of the literature to provide a clear guidance on the state-of-the-art and potential future direction has been lacking. Thus, this review paper is focused on providing up-to-date information on studies carried out so far in the area of fruit detection and localization in specialty crops for robotic harvesting and/or crop-load estimation.

2. Sensors and systems for fruit detection

Accuracy of machine vision systems in the detection and localization of fruit is affected by uncertain and variable lighting conditions in the field environment, variable, and complex canopy

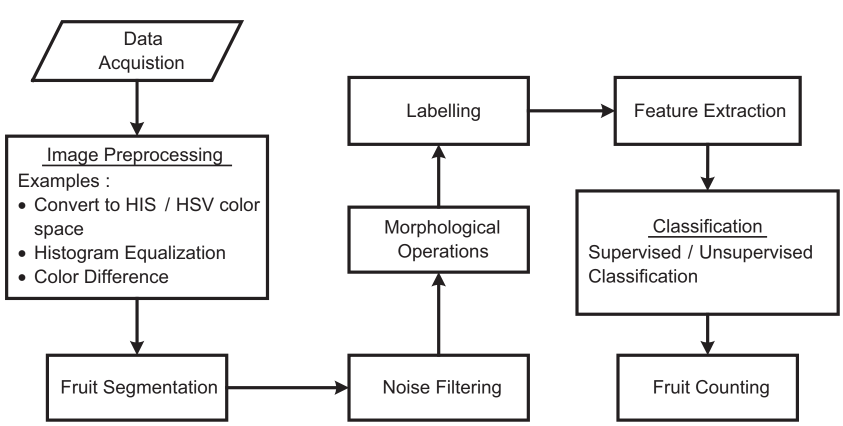

Fig. 1. Block diagram of steps involved in fruit detection.

structures (Karkee and Zhang, 2012) and varying color, shape and size of the fruit. In addition, the accuracy of fruit detection is substantially limited by the occlusion of fruit in canopy images by leaves, branches, and other fruit (Gongal et al., 2015). A number of research projects have been carried out in the past to accurately detect fruit in similar outdoor environments. Methods studied in the past for fruit detection utilized different types of sensors, and various image analysis as well as soft computation methods. The major steps involved in fruit detection (Fig. 1) are; (i) image acquisition; (ii) image preprocessing; (iii) fruit segmentation; (iv) noise filtering; (v) morphological operation; (vi) labelling; and (vii) feature extraction. Also in some studies, classifications have been carried out after feature extraction to improve the detection accuracy. Various types of sensors and image segmentation and classification methods used for fruit detection are reviewed in the following sub-section and summarized in Tables 1–3.

2.1. Black and White (B/W) cameras

B/W cameras were used in some of the earliest studies in detecting fruit based on geometric features (Whittaker et al., 1987). A prototype self-propelled robot developed for fruit harvesting by Grand D'Esnon et al. (1987) under MAGALI project used a B/W camera. Cardenas-Weber et al. (1991) and Dobrusin et al. (1992) also used B/W camera to detect melons based on fruit geometry and texture. Pla et al. (1993) used B/W camera for green immature orange detection based on convex surfaces of the fruit and obtained 75.0% accuracy. On the other hand, Edan et al. (2000) used B/W camera to detect melons based on reflectance, geometric and texture features and reported an accuracy of 82.0–88.0%. The author recognized that use of multiple sensors and combining their information could improve the results. Sites (1988) also used a B/W camera with a color filter (630–670 nm) to amplify the contrast between fruit and background. The author reported 90.0% classification accuracy during nighttime and 84.0% accuracy in daytime with 20.0% false detection. The European Eureka Project (Juste et al., 1991) also used B/W camera in conjunction with red (630 nm), and green (560 nm) color filters and reported an accuracy of 80.0% with 10.0% false detection. However, the B/W camera was replaced later with color camera and an accuracy of 90.0% on fruit detection was reported with a 5.0% false detection. The major disadvantage of B/W cameras was that color information was not available, which is one of the

Table 1
Summary of sensors used for fruit detection.

Sensors	Type	Advantages	Limitations	References
B/W camera	Passive sensor	Minimal effect of variable lighting conditions	Lack of color information	Whittaker et al. (1987), Juste et al. (1991), Edan et al. (2000)
Color camera	Passive sensor	Provides color, geometric and texture information	Affected by lighting conditions	Bulanon et al. (2002), Cohen et al. (2011), Silwal et al. (2014)
Spectral camera	Passive sensor	Can have both spectral and color information	Time consuming	Safren et al. (2007), Okamoto and Lee (2009), Alchanatis et al. (2007)
Thermal camera	Passive sensor	Independent of fruit color	Affected by size of fruit; narrow range of operations during day time	Stajnko et al. (2004), Wachs et al. (2009), Bulanon et al. (2008)

Table 2
Summary of different features used in fruit detection.

Features	Sensors	Accuracy (%)	Limitation	Crops applied	References
Color	Color camera	80–85	Variable Lighting Conditions	Apples, Citrus	Bulanon et al. (2002), Zhou et al. (2012), Qiang et al. (2014)
Geometric	B/W camera	68–75	Occlusions	Tomatoes, Citrus	Pla et al. (1993), Whittaker et al. (1987)
Texture	Color camera	75.3–85	Variable Lighting Conditions; Occlusions; Variable fruit size	Citrus, Green Apple, Pineapples, Bitter Melon	Rakun et al. (2011), Zhao et al. (2005), Kurtulmus et al. (2011), Chaivivatrakul and Dailey (2014)
Integration of features	Color camera, laser range finder, hyperspectral camera	87–90	Variable Lighting Condition, Occlusion, Clustering	Orange, Apple, Mango	Linker et al. (2012), Stajnko et al. (2009), Dobrusin et al. (1993), Payne et al. (2014), Cakir et al. (2013)

Table 3
Summary of classification methods used in fruit detection.

Methods	Sensors	Accuracy (%)	Limitation	Crops applied	References
K mean Clustering	Color camera, thermal camera	53–80	Variable Lighting Conditions	Apple	Wachs et al. (2010), Bulanon et al. (2004)
KNN clustering	Color camera	85–90	Variable Lighting Conditions; Variable Fruit Size	Apple, Banana, Lemon, Peaches, Strawberry,	Linker et al. (2012), Seng and Mirisaee (2009), Kurtulmus et al. (2014)
Bayesian Classifier	Color camera	75	Variable Lighting Conditions	Citrus, Apple, Peaches	Slaughter and Harrell (1989), Chinchuluun et al. (2007), Kurtulmus et al. (2014)
Artificial Neural Network	Color camera	61–84	Variable Lighting Conditions; Occlusion	Apple, Citrus, Peaches,	Plebe and Grasso (2001), Regunathan and Lee (2005), Kurtulmus et al. (2014)
Support Vector Machine	Color camera	92–93	Variable Lighting Conditions; Occlusion; Clustering	Apple, Citrus, Peaches	Ji et al. (2012), Wang et al. (2009), Qiang et al. (2014)

most prominent features of fruit. Thus, limited feature information of fruit makes it difficult to achieve desired level of accuracy on fruit identification using B/W camera.

2.2. Color cameras

Color cameras with CCD (Charged Coupled Device) or CMOS (Complementary Metal-Oxide-Semiconductor) sensor have been widely used by researchers working in machine vision systems for robotic and automated agricultural operations (Slaughter and Harrell, 1989; Linker et al., 2012; Tabb et al., 2006; Bulanon et al., 2002; Cohen et al., 2011; Silwal et al., 2014). Images acquired with color cameras provide opportunity for simple color based segmentation of fruit in addition to geometric, and texture information. Slaughter and Harrell (1989) used color camera to acquire images and detect oranges based on 'HS' components of Hue, Saturation, Intensity (HSI) color space and Red, Green, Blue channels of RGB color spaces of objects and achieved an classification accuracy of 75%. Similarly, Kurtulmus et al. (2014) used color camera for detection of peaches using statistical classifier and neural network. The major disadvantage of this sensor is that the images captured are sensitive to variation in lighting conditions. In addition, detecting apple varieties with less distinct color such as Granny Smith (green colored fruit) will be challenging as the rest of the plant materials, including leaves and branches may have a similar color signature.

2.3. Spectral cameras

Researchers used spectral imaging to detect objects of interest based on their reflectance at different wavelengths. Spectral camera provides spectral information along with the spatial information of objects. Spectral imaging has the potential to detect fruit even when fruit and background color is similar (Kondo et al., 1996; Van Henten et al., 2002; Kane and Lee, 2006; Bulanon et al., 2010; Wang et al., 2013). Multispectral imaging was carried out by Kane and Lee (2007) using a near infrared (NIR) camera and capturing images at three optical bands (1064, 1150, and 1572 nm). Index calculation was done based on images at three different wavelengths, and then Otsu's threshold (Gonzalez et al., 2010) and morphological operations were carried out for fruit segmentation. The author reported coefficient of regression (R^2) of 0.74 between numbers of citrus detected and actual number of citrus in image. The major challenge in the experiment was target shifting due to wind, and change in shadow position between imaging with different filters. The author recommended capturing multiple waveband images at the same time, so that there is no target shift. They also suggested using smart image processing techniques for better detection accuracy. Previous researchers have also used spectral signature including color information from hyperspectral images for fruit detection (Safren et al., 2007; Okamoto and Lee, 2009; Alchanatis et al., 2007). Safren et al. (2007) used multistage algorithm to analyze hyperspectral data in visible and near infrared (NIR) range for apple segmentation. The high dimensional data from hyperspectral camera was first reduced using principle component analysis (PCA). Then homogeneous objects were extracted and classified followed by morphological operations, watershed analysis, and blob analysis. Integration of these methods resulted in a fruit segmentation accuracy of 88.1%. The author recognized that accuracy of hyperspectral imaging is better than multispectral imaging (Kim and Reid, 2004) for fruit detection as it has more information about the reflectance characteristics in the scene. However, the author recognized that feasibility of hyperspectral imaging is limited by time-consuming data acquisition and image analysis process.

2.4. Thermal cameras

Thermal cameras capture temperature signature of objects, which may be helpful in differentiating fruit and background. Researchers (Bulanon et al., 2008, 2009; Stajnko et al., 2004) have used thermal imaging to detect fruit based on the temperature difference between the fruit and background. The fruit absorbs more heat and radiates more heat in comparison with leaves and other parts of the plant canopy, which allows for distinction between those plant materials with thermal imaging (Stajnko et al., 2004; Wachs et al., 2009). Stajnko et al. (2004) used thermal camera for green fruit detection as the detection of green fruit would be challenging with the use of only color information. The pixel value in thermal image was first transformed to RGB color space and then to the chromaticity coordinates. Fruit segmentation was then performed using global threshold and normalized difference index (NDI) (Fig. 2). The maximum regression coefficient between actual and detected apple count was 0.88 for mature fruit. The accuracy of this method was limited by fruit size and exposure to direct sunlight. The accuracy in detecting fruit that was shadowed and deep in the tree canopy was low, as there was not a significant

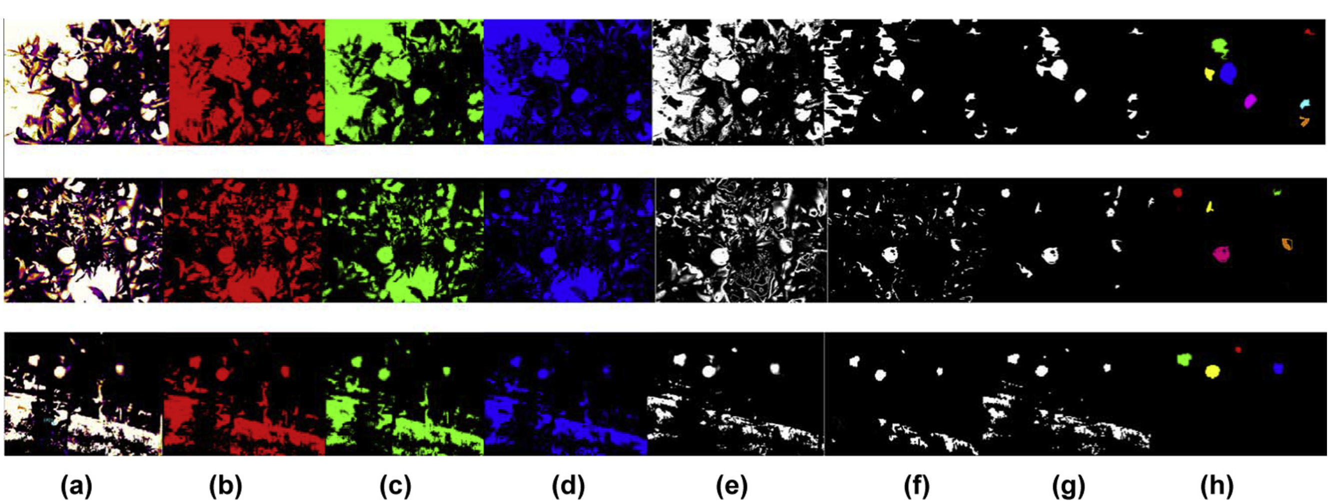

Fig. 2. Results of image processing algorithm used for fruit detection; (a) RGB image; (b) Red image; (c) Green image; (d) Blue image; (e) NDI image; (f) binary image after using NDI index for threshold; (g) fruit and leaves segmentation; and (h) results of fruit detection (Adoption from Stajnko et al., 2004). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

temperature difference between fruit and the background in those areas. The authors recommended that the addition of a shape feature in the fruit detection method could improve the accuracy of the system through identification of partially hidden fruit.

3. Features and methods used in fruit detection

3.1. Color features

Color is one of the most important features used in machine vision system to distinguish fruit from leaves, branches and other background objects in the orchard environment. A number of researchers (Annamalai and Lee, 2003; Qiang et al., 2014; Hannan and Burks, 2007; Slaughter and Harrell, 1989; Bin et al., 2010) have used color-based segmentation in detecting fruit with distinct colors including tomatoes, red apples, peaches, mangoes, pineapples, and citrus. These approaches have achieved accuracy up to 88.0% when only color features were used for fruit detection. However, accuracy of fruit detection based on color is affected by variation in fruit color due to its maturity level, fruit variety, uncertain and varying background features and variable lighting conditions. For example, depending upon the maturity level and fruit exposure to sun over the growing season, Jazz apples in a tree may vary from green, reddish green to red, which makes it challenging to develop fruit detection methods that work robustly on all those apples. In addition, fruit under direct sunlight and shadow during imaging will have different light intensity in the images, and thus would be challenging to detect them.

Bulanon et al. (2002) used luminance and RGB color difference to segment the apples in different lighting conditions. Apple detection was carried out based on the red color difference between the objects and an accuracy of 88.0% was achieved under controlled lighting conditions. The author recognized that accuracy of fruit detection depends on lighting conditions during imaging, and high accuracy of apple detection can be obtained under controlled lighting conditions. Zhou et al. (2012) used color features in both RGB and HSI space for apple detection (Fig. 3). The difference between red, green, and blue color channels was used as a measure to segment red and green apples from the background. Further, a threshold value in saturation channel was also used to segment red apples. The regression coefficient between automated and manual apple count was 0.80 and 0.85 for green immature and red mature apples. The accuracy of this method was affected by occlusion of fruit by leaves, branches, and clustering of fruit. Thus, for better accuracy fruit detection should be carried out using multiple features such as color, shape, texture, and reflectance to overcome challenges like clustering, occlusion, and variable lighting conditions.

3.2. Geometric features

Geometric measures such as shape and size provide another set of distinct features of fruit such as citrus, apple, and pear. They are more rounds in shape than branches and leaves, and have genetic size constraints. Detection of green fruit (immature fruit or green variety) including apples, citrus, and pears, whose color signatures are similar to image background (e.g. leaves and branches), is easier with geometric features such as shape. These features are also, in general, invariant to lighting condition, which makes it appropriate for its application in orchard environment. However, the main problem with this method is the occlusion of fruit by leaves, branches and other fruit, which results in change in shape, size, and other geometric parameters of fruit.

Whittaker et al. (1987) used shape (circular) feature to identify tomatoes in the B/W images using Circular Hough Transform (CHT). CHT is used to identify circles or part of circles in an image in which tomatoes that are partially occluded by leaves and background can still be detected. The accuracy on correctly classifying tomatoes was 68% with 42% false detection as tomatoes in this study. The authors recommended using CHT with gray level contrast as some leaves were also detected as fruit in their study because of the circular contour of the leaves. The authors also recognized that the algorithm used was computational intensive, which may limit its application in real-time environment.

3.3. Texture features

Textures are another distinguishable feature of fruit that can be helpful in separating fruit from the background. Fruit generally has smoother surfaces than background objects such as leaves and stems. During the detection process based on texture features, the surfaces having the same texture are isolated and then edges of the isolated surface are detected (Zhao et al., 2005). The surface color does not affect the texture analysis so it can be used for the detection of fruit with similar color to leaves and stems.

Numerous studies have used texture analysis in fruit detection methods (Rakun et al., 2011; Zhao et al., 2005; Kurtulmus et al., 2011; Chaivivatrakul and Dailey, 2014). Kurtulmus et al. (2011) used Gabor texture analysis proposed by Zhang et al. (2002) to segment green citrus from the background. The result from texture

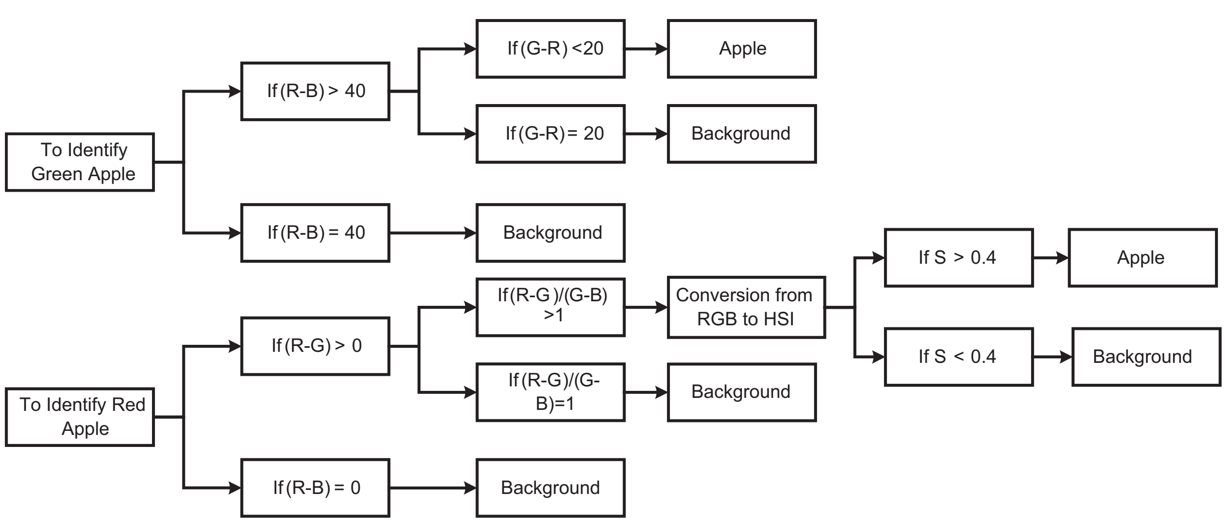

Fig. 3. Flowchart explaining different steps of fruit segmentation (Adopted from Zhou et al., 2012).

analysis was integrated with blob analysis, and the 'Eigen Fruit' approach with 75.3% of fruit identified. There was 27.3% false detection of leaves and stems as green citrus. The authors recognized that variable lighting condition, visual complexity of the background and varying fruit size limited the accuracy of the system. Zhao et al., 2005 also mentioned that limitation of texture-based segmentation in orchards is variable lighting conditions as it affects the texture properties of fruit.

3.4. Integration of color, geometric and texture features

The variable lighting condition in orchards will affect the intensity of reflected light whereas partial occlusion of fruit by leaves, branches, and clustering of fruit will affect the geometric features of fruit in images. In such scenario, detection based on a single group of features (e.g. color, geometry, or texture) may not be the best approach. Researchers (Linker et al., 2012; Stajnko et al., 2009, 2007; Dobrusin et al., 1993; Payne et al., 2013, 2014; Plebe and Grasso, 2001; Cakir et al., 2013) have integrated multiple features from two or more feature groups to improve the accuracy and robustness of fruit detection methods. Hannan et al. (2010) combined color, shape, and size features to segment oranges in tree canopy images. Chromaticity in 'r' $\left(r = \frac{R+G+B}{R}\right)$ was used for fruit segmentation where area filtering was applied to remove small noises in the image. Then, perimeter based shape detection was conducted to detect oranges in clusters. The authors reported that 90.0% of visible oranges in an image were detected with a false detection rate of 4.0%. They also recognized that variable lighting condition is the major challenge in this experiment. Although the accuracy of fruit detection is high, it did not address the partially occluded fruit. Thus, the accuracy on total fruit detection on a tree will be lower than 90.0%, as not all fruit in the tree are completely visible in the image.

Jiménez et al. (1999b) used a laser range finder to acquire range and attenuation data and combined the information with color and geometric features to identify orange and green oranges. Any pixel beyond the range of expected distance to fruit was filtered out as noise. The accuracy on identifying visible oranges was 58.0% based on color information, and 74.0% based on shape information. The accuracy increases to 87.0% with no false detection when both shape and color features were integrated. Green fruit in the images affected the accuracy of this system. Thus, authors recommended improving the shape-based recognition method using profile shape and curvature information along with the contour shape information.

4. Image classification methods for fruit detection

4.1. Unsupervised classification: K-means clustering

One of the widely used unsupervised classification methods is K-means clustering. K-means is a learning algorithm, which separates the input data set into different clusters based on their inherent distance between each other. It minimizes the sum of distance between the objects and respective cluster centers. It is an iterative process, which moves objects between the clusters until the sum is no longer minimized (Shapiro and Stockman, 2001).

Wachs et al. (2010) used thermal images and color images for green apple detection. K-means clustering based on 'a' and 'b' channels from $L*a*b$ color space (Hunter, 1948) was carried out to segment apples from both images. Morphological operation and CHT was then carried out to remove noises and improve the classification accuracy. The accuracy on detecting visible fruit was 38.8% in color images, 50.6% in thermal images, and 53.2% when information from both thermal and color images was used. The author recognized that variable lighting conditions and relying heavily on color information were the major factors for lower accuracy in this system. Bulanon et al. (2004) also used K-means clustering for red apple detection. RGB color space was transformed to chromacity 'rgb' ($r = \frac{R}{R+G+B}, g = \frac{G}{R+G+B}, b = \frac{B}{R+G+B}$) color space for fruit segmentation followed by low pass filtering to remove noises in the image. The authors reported that 80.0% of fruit pixels of visible apples were correctly classified with 3.0% error rate.

4.2. Supervised classification

4.2.1. Bayesian classifier

Bayesian classification is a probabilistic classifier based on Bayes' theorem, which make statistical interpretation based on prior knowledge and probability distributions. The Bayesian classifier rule is based on minimization of Bayes risk, minimization of probability error, or maximization of posteriori probability depending on the priori probability (Duda et al., 2012).

Slaughter and Harrell (1989) used a Bayesian classifier to classify oranges based on color information. The mean and covariance of red, green, blue (RGB) values of fruit and background pixels were determined and then priori probability was assumed. Thirteen images captured at different lighting conditions, different backgrounds (with only leaves, or with bright background, or with varieties of natural background), and with different occlusion levels were used for evaluation. The authors reported an accuracy of 75.0% on classifying orange pixels; false positive and false negative were not reported in the study. Variable lighting conditions, intense background lights, and occlusion of fruit were the major reasons for the inaccuracy in this study. Chinchuluun et al. (2007) also used Bayesian classifier for citrus image segmentation based on 'S' and 'I' components of Hue, Saturation, Value (HSV) color space and luminance (Y) and chrominance (IQ) (YIQ) color space respectively. The algorithm was evaluated with images of fruit harvested in field. The authors reported a regression coefficient of 0.89 between manually and automatically counted fruit, the actual accuracy of the system was not reported. The authors recognized that the error was mainly due to variability in lighting condition, which causes uncertainty in priori probability. The major disadvantage of Bayesian classifier is that it needs prior probabilities information from training image (Bulanon et al., 2002), which may not be available in outdoor application. Chinchuluun et al. (2007) also reported that in case of variable lighting condition the pixel color value of fruit changes, which affects the prior probabilities and hence the performance of Bayesian classification.

4.2.2. KNN clustering

KNN clustering is supervised learning method, used widely in classification and regression studies. It is used to classify unknown feature vector to the class that have most common properties among its K nearest neighbor using existing training samples. Majority voting of a K number of nearest neighbor are used for classification (Shapiro and Stockman, 2001). Various researchers including Linker et al. (2012), Seng and Mirisaee (2009) and Kurtulmus et al. (2014) have used KNN clustering classification to classify different fruit.

Linker et al. (2012) used KNN classifier for classification of green apples. The results were further improved by apple surface detection using seed area detection and growing region based on color and texture information, then contour of each region is segmented into an arc so that the apple can be recognized. The authors reported that 85.0% of green apples were detected under direct lighting conditions. They also claimed that 95.0% of green apples were detected under diffused lighting conditions. The authors recognized that variable lighting conditions and fruit size are the main factors reducing the accuracy of the system. Seng and Mirisaee (2009) also used KNN for identification of apples, bananas, lemons, and strawberries. Fruit color, shape, area, and perimeter features were used to distinguish the fruit from each other. The authors reported an accuracy of 90.0% on recognition of fruit. False positive and false negative errors were not reported in the study.

4.3. Soft computing methods

Soft computing methods are often used to perform modeling and analysis of complex problems, and to provide solutions, which are tolerant to imprecision, uncertainty, partial truth and approximation (Huang et al., 2010). Soft computing methods learn from experimental data and give output of unseen input by approximation and interpolation. Soft computing methods have long been investigated for the application in scientific research and engineering computing. In the following paragraphs, some of the major soft computing methods used in fruit identification are discussed.

4.3.1. Artificial Neural Network (ANN)

An ANN is a supervised learning algorithm that has the ability to learn from the environment through an iterative training process and improve its performance after each iteration. ANNs have interconnections between neurons that feed the information from one part of network to other to compute results from given input. The network consists of multiple layers of neurons, which depend on the complexity of the system. The more complex the system, the more layers of input or output neurons are found (Dahikar and Rode, 2014). Plebe and Grasso (2001) used neural network to identify oranges based on color features. The network was trained by back propagation method (Rumelhart et al., 1986; Benhanan et al., 1992) where neurons sent the signal forward and errors were propagated backward, using samples of pixels of oranges and backgrounds from a set of 800 images captured in different lighting conditions. The authors reported that 87.0% of oranges were detected in the images with 15.0% false positive and 5.0% false negative. The author recognized that variable lighting conditions and fruit occlusion were the major causes of error in their study.

Regunathan and Lee (2005) used multi-layer, error backpropagation neural network for citrus detection based on hue, luminance, and saturation (HLS) color space. H and S values of every pixel were used to classify the pixels in citrus. Fruit was then segmented and morphological operation and watershed transformation was applied. The author reported mean percentage error of 39.6% on detecting fruit visible in the image. False positive and false negative errors on detecting fruit were not reported in the study. Wachs et al. (2010) used multi-layer neural network trained with back propagation algorithm to detect citrus in color and thermal images based on thermal value and color features. The pixel value of fruit and background in $L*a*b$, HSV, and RGB color spaces were stored, and then each classifier was trained and tested for each color space. The accuracy of fruit detection in the color image was improved by combining the outputs of all three classifiers using majority voting. The authors reported an accuracy of 66.3% on detecting visible fruit in the image with 33.7% false negative. The results from color image and thermal image were also combined for better identification accuracy reaching 74.4% with 25.6% false negative. The authors recognized that accuracy could further be improved by increasing the sample size and using morphological information in addition to color information.

4.3.2. Support Vector Machine (SVM)

SVM is also a supervised statistical learning algorithm that has been used for linear and non-linear regression analysis, and pattern classification. For linearly separable classification, SVM separates the two classes with a maximum margin between them by a hyper-linear plane. Similarly, in nonlinear separable classification, feature vectors are mapped to a new feature space that is linearly separable and then the classification is done based on linear SVM separation (Wang et al., 2009).

Ji et al. (2012) and Wang et al. (2009) used SVM based on pixel value of apple in hue, luminance, and saturation (HLS) color space and shape feature for red apple detection. The color features and shape features of apples were extracted from 150 sample images. Classification was then carried out with three different SVM kernel functions (Poly, Radial basis function (RBF), and Sigmoid) based on color, and shape as well as integrated color-shape features. The fruit detection accuracy was 93.0% when RBF kernel function was implemented based on integrated color and shape features. The authors did not report false positive and false negative errors and or if the accuracy is based on only visible apples or total apples in the images.

Qiang et al. (2014) also used multi-class SVM classification method for detection of citrus fruit. RBF kernel function was used to classify citrus, leaves, and branches based on RGB color features. Citrus detection accuracy of 92.4% was achieved based on a manual count in the images, however false positive and false negative errors were not reported. They recognized that variable lighting conditions, occlusion, and immature fruit were the major factors reducing the accuracy of the system.

Sengupta and Lee (2012) first used CHT to detect spherical objects in the image followed by a SVM classifier based on local texture and Tamura texture features (Tamura et al., 1978) to classify green citrus and background. To reduce false positive classification, scale invariant feature transform (SIFT) algorithm (Lowe, 1999, 2004) was applied to identify local features of fruit that were invariant to translation, scaling, rotation or occlusions. Accuracy of 81.7% in fruit detection with visibility of more than 50.0% of fruit's area with 25.6% false positive were reported. It was recognized that the variable lighting conditions, partial occlusions, and green fruit color are the major factors that cause inaccuracies in the detection. As discussed briefly in various places earlier in the paper and as summarized in Table 3, there are various challenges in fruit detection for agricultural applications including crop load estimation and robotic harvesting. More discussion on those challenges and potential future direction for research and development will be presented in Section 6.

5. Sensors and systems for fruit localization

Fruit localization in trees is another important part of machine vision system for robotic harvesting and sizing of fruit for crop-load estimation. The major challenges in localizing fruit are the displacement of fruit by wind or other factors during imaging, and occlusion of fruit (Sarig, 1993). Numerous studies have been carried out to locate fruit in tree canopies with reasonable successes (Harrell et al., 1989, 1990; Slaughter and Harrell, 1989; Whittaker et al., 1987; Kondo and Kawamura, 1983; Mehta and Burks, 2014; Font et al., 2014). The sensors and methods used in the past for fruit localization in the agricultural environment are critically reviewed in the following sub-sections.

5.1. Single camera

Color cameras, consisting of either CCD sensors or CMOS sensors have been used for localizing fruit and for tracking the trajectory of harvesting robots. Baeten et al. (2008) used a CMOS color camera for apples localization, and for guiding a robot for harvesting. The relationship between the focal length of the camera, pixel

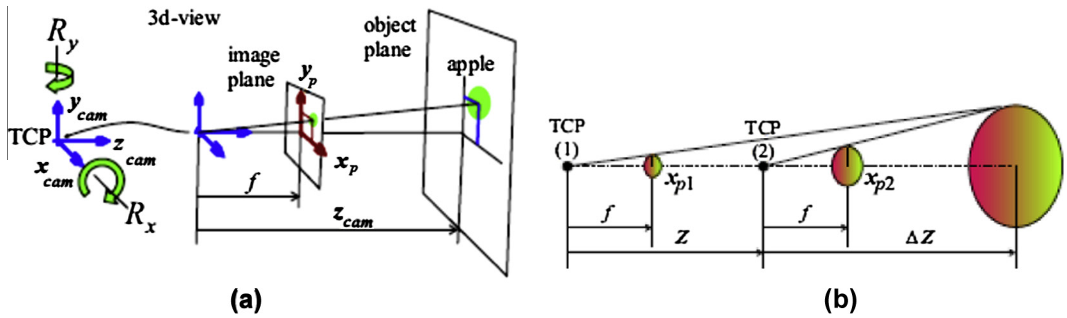

Fig. 4. (a) 3D view of the camera model with respect to camera frame; and (b) Illustration of calculation of remaining distance from camera to the object (Baeten et al., 2008).

size, and center of the apples in the image plane was used to calculate the distance between the camera and the fruit. Several images were taken to calculate the remaining distance to the apples by triangulations (Fig. 4) while approaching the apples.

Zhao et al. (2011) also used a CCD camera on an end-effector of a harvesting robot to localize fruit in the image plane. The centroid of the fruit in the image plane was determined, and then the manipulator arms were rotated to the corresponding image center and fruit center together. Then the flexible arm was spread until the fruit entered the gripper. The accuracy of the system in localizing the fruit and tracking the trajectory was not reported in the paper.

Parish and Goksel (1977) also used a single camera to determine the centroid of fruit in the image and performed a mathematical transformation to relate 2D coordinates of objects to a 3D coordinate system. A touch sensor was used to determine whether the end-effector reached the apples. Peterson et al. (1999) also used a camera along with an encoder for tracking the trajectory of a fruit-harvesting robot. The position of a rapid displacement actuator (RDA) in *X* and *Y*-axis was tracked by the encoder which were then converted to pixel coordinates. The accuracy obtained on moving RDA to required position by reading the position from encoder was 90.0%. Similarly, Mehta and Burks (2014) used a single camera for depth estimation of citrus from the robot base. Euclidean position of fruit was obtained using pixel coordinates of fruit, size of fruit and intrinsic parameters of the camera. Then the position of fruit with respect to the robot base was obtained by transforming the coordinates from camera to robot base reference frame using extrinsic camera parameters. The accuracy on estimating the position of citrus was approximately 15 mm in their study. The authors argued that the use of depth estimation using a single color camera could be a better solution than using computationally complex stereo vision or a triangulation scheme. It was recognized that the error in depth estimation was partly caused due to error in identifying fruit's centroid and size, as the depth estimation is dependent on fruit size in this study. However, this method requires a moving sensor to estimate fruit location continuously, which can potentially reduce the speed at which the manipulator and end-effector can be moved to desired location.

5.2. Laser range finder

Laser range finders operate on the principle of Time-of-Flight (TOF) of light. A laser range sensing unit consists of a laser source to emit pulsed laser beams and a sensor to receive the beam reflected back from the objects. The time taken by the laser beam to reflect back from the objects is proportional to the distance between the object and the sensor (Feng et al., 2012). Laser range finders can provide range information of an entire scene using a scanning sensor. The scene can be scanned horizontally and vertically to create 3D coordinate map of the scene. Researchers have widely used laser range finders in agricultural automation and robotics including fruit localization in tree canopies (Ceres et al., 1998; Jiménez et al., 1999a; Feng et al., 2012).

Bulanon et al. (2010) used a laser sensor to determine the distance to fruit in apple orchards. A color camera and a laser sensor were mounted together in a cylindrical manipulator, which was controlled by a visual servo method (Bulanon et al., 2005). The target fruit center was aligned in the center of image by visual servoing and then a laser range sensor measured the distance to the fruit. The accuracy of this system in estimating distance to fruit was reported to be ±3 mm. However, this system is bulky and slow which might increase complexity during data acquisition.

5.3. Stereovision

Stereovision systems consist of two or more cameras separated by a certain distance. Multiple images captured by individual cameras are matched together to estimate spatial displacement (or disparity) of the object in two images. Image disparity is then converted to distance to objects from the camera using relative camera locations and orientations, and focal lengths of the cameras (Shapiro and Stockman, 2001). Researchers have widely used this method to simultaneously identify and locate fruit for harvesting or crop-load estimation (Font et al., 2014; Tanigaki et al., 2008; Kitamura and Oka, 2005).

Plebe and Grasso (2001) installed stereo cameras in two arms of an orange harvesting robot. Stereo matching of orange centers was performed to locate oranges in a 3D coordinate system. An ANN was applied to perform stereo matching of two images captured by two cameras installed in two different arms. The accuracy of the system to reach the target was affected by wind and variable lighting conditions, along with the efficiency of the hardware component of the system. Actual localization accuracy of the system was not reported. Wang et al. (2013) also used a stereovision camera for localizing fruit in the global coordinates to identify repeated counting of apples due to multiple imaging. The apples that were at a distance less than twice the diameter were considered as repeated occurrences of the same apple in multiple images. This study did not report accuracy of fruit localization, or false negative and positive in identifying repeated apples. The authors recognized that a large distance threshold was used to detect repeated apples in multiple images to compensate for the error due to stereo triangulation. Hannan and Burks (2004) also reported that the practical application of stereovision system is limited due to its complexity and long computation time.

5.4. Time-of-Flight (TOF) of light-based 3D cameras

TOF cameras (also called 3D cameras) use a similar principle as a laser range finder, but create a distance map of the entire scene at once. The Photonic Mixing Device (PMD) CamCube 3.0 (PMD Technologies, Siegen, Germany) is an example of such a camera. In addition to mapping distance, these cameras provide intensity and 3D coordinates of objects in the field of view. Application of TOF camera in agricultural automation has been investigated for 3D reconstruction of apple trees (Karkee et al., 2014), determination of interplant spacing (Nakarmi and Tang, 2012) as well as for localization of fruit in trees (Gongal et al., 2015).

Gongal et al. (2015) used a TOF camera to determine 3D coordinates of fruit in apple trees trained to a narrow fruiting wall in Washington State. The 3D coordinate information was used to identify repeated apples that were visible in images captured from two opposite sides of the tree canopy. The method achieved an accuracy of 87.0% in identifying repeated apples. The work did not assess the absolute location accuracy with this camera system but observed that some of the errors in the system were due to error in registering 3D location information with the apples identified in the color camera images.

Data acquisition and processing for 3D localization can be faster with 3D camera compared to other 3D sensing systems such as a stereovision camera. These cameras are also claimed to provide better accuracy in 3D reconstruction compared to stereovision systems (Beder et al., 2007). However, TOF cameras available in the market have lower resolution and are expensive.

As discussed in this section, various techniques have been investigated in the past for 3D localization of fruit in field environment, each with its own pros and cons. However, it was difficult to find studies reporting absolute accuracy of each technique and comparisons of performance between those techniques under the same outdoor environment, which could be an area of investigation in the near future. Various challenges in 3D localization of fruit and potential direction for future research and development are further discussed in the following section.

6. Challenges and future direction

Further research and development is necessary to establish the technical as well as commercial viability of apple detection and localization systems for fruit harvesting and/or crop-load estimation. Most of the studies suggested that occlusions, clustering, and variable lighting conditions of fruit were the major challenges limiting the fruit detection as well as localization accuracies in orchard environment. These challenges need to be further addressed through integrated horticultural and engineering approaches for improved image segmentation, and for increased overall accuracies and robustness of fruit detection and localization systems.

6.1. Fruit detection

First, a systematic method of various horticultural operations including tree training, pruning, pollination, and thinning has the potential to present fruit to machine vision system with improved visibility, and reduced occlusion and clustering. Such effort will simplify the task of machine vision system and will improve the fruit detection accuracy and speed, and simplifies navigation of robot hands during harvesting.

To address the issue of variable lighting conditions, one of the potential areas of research and development could be to provide a controlled lighting environment for imaging in agricultural fields. To create controlled lighting environments for imaging, Wang et al. (2013) and Payne et al. (2014) used nighttime imaging. However, this technique limits the operation time of a machine in commercial application. Silwal et al. (2014) and Gongal (2014) reported that the use of over-the-row platform with a tunnel structure could create controlled lighting conditions in the orchard environment that can be used during both daytime and nighttime.

Some literatures have suggested that fruit detachment efficiency can be improved if fruit is rotated or twisted in particular ways relative to orientation of fruit and stem (Bulanon and Katoka, 2010; Tong et al., 2014). Further evaluation of manual fruit picking process specific to different tree fruit crops and understanding of fruit picking dynamics including forces and motion involved during manual picking is essential for developing effective and efficient fruit picking end-effectors. Consequently, future research in machine vision system should also focus on estimating orientation of fruit, and orientation and location of stem. Direct detection of fruit stem with a machine vision system is challenging as it is small and often blocked by the fruit. However, geometric parameters of the fruit such as location of calyx, outline or bounding box of fruit, and the symmetry of the shape could provide a way to estimate fruit orientation, and stem location and orientation. Investigation of deep learning models used in pose estimation (Yu et al., 2013; Liang et al., 2014) for application in outdoor environment may be another area of research with potential contribution to estimating fruit and stem orientation.

The problem becomes even more challenging for fruit that are only partially visible, which is prevalent even in the modern fruiting wall orchards (Silwal et al., 2014). When only a portion of fruit is detected, estimating its shape and therefore the orientation of fruit and stem is difficult. To address this challenge, reconstruction of whole fruit using the detected portion is essential. Image mining based on matching of partial fruit outlines, and first and second order derivatives of outlines could be one of the ways to reconstruct fruit from the detected portions. Such reconstruction will also help accurately size apples and improve crop-load estimation.

Understanding growth patterns of various types of fruit and different cultivars of a given crop is another area of research that affects the potential of machine vision system. For example, stem length and fruit shape may vary substantially between various apple cultivars. If a good understanding of such parameters is available, machine vision system can use such knowledge to improve its effectiveness in detection and/or reconstruction of partially visible fruit. In addition, future research and development in machine vision system should focus on detecting feature points in the apple body so that fruit could be grabbed in desired positions for effective and efficient picking.

Human machine collaboration will be increasingly essential in achieving desired level of accuracy and speed in all the functionalities of a machine vision system. Even with the most advanced fruiting wall tree architectures and future improvement in horticultural approaches in developing more machine-friendly orchards, some fruit may still be behind leaves and under shadows making detection task difficult, time consuming or both. However, if some level of support is provided by human operator to detect fruit and stem position and orientation in difficult canopy areas and correct the errors made by the machine, overall speed and accuracy of the system may potentially be improved. It is also worth noting that machine vision task for fruit harvesting is somewhat simplified by the fact that removal of one fruit can improve exposure of other fruit. If image acquisition and fruit detection is repeated frequently during harvesting process, the overall visibility of fruit in the canopies may improve, particularly when there are occlusions due to other fruit and/or fruit clustering. Though the general outcome of human machine collaboration and repeated fruit detection is obvious, quantitative results on what level of accuracy and speed is possible with such a system is yet to be determined, particularly for newer tree fruit orchards.

6.2. Fruit localization

Fruit grasping rate of a robotic harvesting system depends on fruit localization accuracy and on the mechanical design of the end-effector and its control system. Although end- effector design can tolerate certain level of location error while grasping the fruit, larger tolerance may lead to grabbing multiple fruit or unsuccessful grab when fruit are in cluster or are in close vicinity of other objects such as leaves and branches. The desired level of 3D localization accuracy for crop-load estimation depends on the size of fruit when it comes to avoiding duplicate counting. When sizing is concerned, higher location accuracy will generally lead to higher accuracy in fruit size estimation, though the effect may be minimal as long as the location accuracy is sufficient for counting.

Some of the major issues causing inaccuracies in fruit localization include errors in fruit detection, as well as fruit occlusion, and inaccuracies of distance measurement sensors/systems. All aspects of future improvement in horticultural and fruit detection systems discussed in Section 6.1 will benefit 3D localization system. If fruit are presented to sensors more clearly and without clusters, there will be less interference with unwanted information/signal, improved image/data processing results and improved 3D location accuracy. In addition, when fruit are clearly visible and are separate from each other, robotic end-effector may be able to tolerate more error on fruit location, essentially reducing the complexity and cost of 3D localization system.

Studies on fruit localization with different types of sensing methods have been carried out to optimize the cost, accuracy, speed and robustness of robotic harvesting system (Section 5). Use of single camera is one of the simplest and most cost effective technologies, but distance information needs to be updated continuously during the movement of the harvesting end-effector in this system. In addition to repeated computation of distance and direction to the fruit, such system will also require repeated computation of inverse kinematics of the robotic manipulator and frequent updates of the control inputs to actuators potentially slowing down the overall manipulator speed. If 3D location can be estimated with sufficient level of accuracy at the beginning of manipulator movement (which also depends on end-effector design), we may be able to avoid the need of visual servoing and potentially improve fruit reaching speed.

Stereovision camera is another option, which addresses the issues with the use of a single camera-based system but is also relatively complex and accuracy is low, particularly in the outdoor environment where stereo matching is problematic. In general, laser range finder (Bulanon et al., 2010) has achieved better location accuracy (±3 mm) compared to other sensors being used currently. However, this system is comparatively bulky, slow, and costly.

Another direction for future research could be fusion of multiple sensors for fruit detection and localization as one type of sensor could compensate the limitation of other type. Obstacle detection and object tracking using fusion of laser and stereo camera have been investigated in the past with promising results (Labayrade et al., 2005; Klimentjew et al., 2010; Baig et al., 2011). One specific approach to reduce sensor cost in laser-stereo fusion could be to use a grid of small number of visible lasers. The grid of laser signals will then provide distinct features for stereo-matching as well as accurate references for further calibration of stereo-based 3D information. Investigators have also used fusion of 3D cameras with color camera for localization in agricultural fields with potential to provide high accuracy with minimal computational demand (Lindner et al., 2007). Even though the resolution of 3D camera is limited, distance measurements on the available pixels have potential to be more accurate. It is then possible to integrate low-resolution 3D camera images (with higher location accuracy) with high-resolution color images so that resolution of the 3D data points may also be improved through some type of interpolation.

Affordability of 3D sensing systems is one of the most limiting factors in accurate 3D localization. One of the areas for future research could be on utilization of the low cost gaming sensors such as Microsoft Kinect (Microsoft Corporation, Redmond, WA) in detecting and locating fruit in the field environment. As the new version (Ver. 2) of Kinect works on the principle of time-of-flight of light, which is similar to the working principle of the 3D PMD camera, Kinect could potentially be a low cost substitute for high accuracy 3D information. This sensor could also be a good option for fusion with other sensors such as a color camera. However, it is noted that further research is necessary to investigate the ways to improve the performance of this sensor in outdoor environment, including potential use of over-the-row tunnel structure to block bright ambient light.

Conceptually, sensor fusion will increase number of sensors used in the overall system, which then can increase system complexity and cost. More comprehensive research and development in sensor fusion including wider variations of sensors would be essential to optimize the cost, speed and accuracy of 3D fruit localization. A complementary area of investigation to sensor fusion could be ‘under-sensed robotic systems’, a term with comparable meaning to ‘under-actuated robotic manipulator/hand’, where the focus would be to identify ways to avoid or minimize the sensors used in a robotic system. One area of recent research and development is to focus on fruit detection and 3D localization with one or a pair of sensors capturing a wider canopy area and providing location information of fruit to multiple robotic arms together (Scarfe et al., 2009). Researchers in the past have focused on end-effector mounted sensors for accurate detection and/or localization of fruit. With higher resolution and more accurate sensors and improved image/data processing capabilities, there is a new focus on using one sensor to support a large number of robotic arms working in parallel. Such system will potentially reduce the cost and complexity of overall system, and improve manipulation speed of robotic arms, all of which are crucial for the commercial success of robotic fruit harvesting systems.

7. Summary

This paper presented a comprehensive review of the studies carried out in the area of fruit detection and localization in tree canopies. Various types of sensor systems and image processing methods and their accuracy were discussed for fruit detection. The accuracy on fruit detection was reported to be 70.0–92.0% for citrus and apples. Ji et al. (2012) reported the highest accuracy (92.0%) in identifying red apples using a color camera and applying Support Vector Machine (SVM) for image classification. Use of supervised learning methods for fruit classification generally accredited a good performance but they increase the system complexity and require sufficient, accurate training samples to achieve good performance. Fruit detection accuracy was also generally better when different features of fruit such as color, shape, and size were integrated compared to classification based on only one type of features. Stajnko et al. (2009) and Hannan et al. (2010) reported an accuracy of 89% and 90% on detecting apples and oranges respectively based on integrated use of color and geometric features. Numerous investigations have also been reported in the area of 3D localization of fruit and other objects in the agricultural environment. In general, laser range finder (Bulanon et al., 2010) has achieved better location accuracy (±3 mm) compared to other sensors being used currently. However, this system is comparatively bulky, slow, and costly.

We also presented the current challenges and potential future directions in the area of fruit detection and localization for robotic harvesting and/or crop-load estimation. Most of the studies suggested that occlusions, clustering, and variable lighting conditions were the major challenges for the accurate detection and localization of fruit in the field environment. Thus to improve fruit detection and localization accuracy for robotic harvesting and crop-load estimation, further research and development should focus on the ways to address these limitations and improve accuracy, speed and robustness of the sensing system while reducing the overall complexity and cost. Horticultural modifications, sensing platform that can improve uniformity of lighting condition, sensor fusion techniques, and human machine collaboration could be some of the areas for future research and development.

Finally, it is noted that robotic Kiwi harvesting work presented by Scarfe et al. (2009) was found to be one of the most successful applications of robotics for fruit harvesting. They used color camera to identify fruit and estimate fruit size. It was reported that the robotic system achieved a system level harvesting speed of 1 fruit per sec. Even with that level of speed on kiwi harvesting, the industry has not yet seen a commercial adoption of the machine. As discussed before, challenges including limited robustness, insufficient speed, and/or high cost need to be addressed further before such technologies could be widely adopted.

Kiwi is grown in vines with fruits hanging mostly down from the canopy allowing relatively easier access to the fruits (both in terms of locating fruits and reaching them). When it comes to other tree fruit crops such as apples, even the formal fruiting wall architectures will present comparatively more challenging canopy structure for machine vision and robotic picking operation. However, the success of machine vision and overall robotic operation in Kiwi harvesting certainly provides a strong foundation for developing machine-vision based automation and robotic harvesting systems for other tree fruit crops such as apples and pears.

Acknowledgements

This research was supported in part by Washington Tree Fruit Research Commission, and by Washington State University Agricultural Research Center federal formula funds, Project No. WNP0745 and No. WNP0728 received from the U.S. Department of Agriculture National Institutes for Food and Agriculture (NIFA) and Washington State University Irrigated Agriculture Research and Extension Center. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the authors and do not necessarily reflect the view of the U.S. Department of Agriculture, Washington State University, or Washington Tree Fruit Research Commission. Authors would also like to acknowledge Dave Allan, Allan Bros. Inc. and Yakima Valley Orchards, Prosser, WA for their support in field data collection.

References

Alchanatis, V., Safren, O., Levi, O., Ostrovsky, V., Stafford, J.V., 2007. Apple yield mapping using hyperspectral machine vision. In: 6th European Conference on Precision Agriculture, Skiathos, Greece, pp. 555–562.

Annamalai, P., Lee, W.S., 2003. Citrus yield mapping system using machine vision. In: ASABE meeting presentation, Paper Number: 031002, St Joseph, MI.

Baeten, J., Donné, K., Boedrij, S., Beckers, W., Claesen, E., 2008. Autonomous fruit picking machine: a robotic apple harvester. In: Laugier, Siegwart, R., (Eds.), Field and Service Robotics (42), pp. 531–539.

Baig, Q., Aycard, O., Vu, T.D., Fraichard, T., 2011. Fusion between laser and stereo vision data for moving objects tracking in intersection like scenario. In: Intelligent Vehicles Symposium, IEEE (IV), pp. 362–367.

Beder, C., Bartczak, B., Koch, R., 2007. A comparison of PMD-cameras and stereo-vision for the task of surface reconstruction using patchlets. In: Proc. CVPR.

Benhanan, U., Peleg, K., Gutman, P., 1992. Classification of fruits by a Boltzmann perceptron neural network. Automatica 28, 961–968.

Bin, L., Maohua, W., Ning, W., 2010. Development of a real-time fruit recognition system for pineapple harvesting robots. In: ASABE Annual International Meeting, Paper Number, 1009510, St Joseph, MI.

Bulanon, D.M., Burks, T.F., Alchanatis, V., 2008. Study on temporal variation in citrus canopy using thermal imaging for citrus fruit detection. Biosyst. Eng. 101 (2), 161–171.

Bulanon, D.M., Burks, T.F., Alchanatis, V., 2009. Image fusion of visible and thermal images for fruit detection. Biosyst. Eng. 103 (1), 12–22.

Bulanon, D.M., Burks, T.F., Alchanatis, V., 2010. A multispectral imaging analysis for enhancing citrus fruit detection. Environ. Control Biol. 48 (2), 81–91.

Bulanon, D.M., Kataoka, T., Okamoto, H., Hata, S., 2004. Development of a real-time machine vision system for the apple harvesting robot. In: SICE Annual Conference (1), pp. 595–598.

Bulanon, D.M., Kataoka, T., Ota, Y., Hiroma, T., 2002. AE—automation and emerging technologies: a segmentation algorithm for the automatic recognition of Fuji apples at harvest. Biosyst. Eng. 83 (4), 405–412.

Bulanon, D., Kataoka, T., Okamoto, H., Hata, S. 2005. Feedback Control of Manipulator Using Machine Vision for Robotic Apple Harvesting. ASAE Paper Number: 053114. St.Joseph, MI.

Bulanon, D., Katoka, T., 2010. Fruit detection system and an end effector for robotic harvesting of Fuji apples. Agric. Eng. Int.: CIGR J. 12 (1), 203–210.

Cakir, Y., Kirci, M., Gunes, E.O., Ustundag, B.B., 2013. Detection of oranges in outdoor conditions. In: Second International Conference on In Agro-Geoinformatics, pp. 500–503.

Cardenas-Weber, M., Hetzroni, A., Miles, G.E., 1991. Machine vision to locate melons and guide robotic harvesting. ASAE Paper Number: 91–7006, St Joseph, MI.

Ceres, R., Pons, J.L., Jimenez, A.R., Martin, J.M., Calderón, L., 1998. Agribot: A robot for aided fruit harvesting. Ind. Robot 25 (5), 337–346.

Chaivivatrakul, S., Dailey, M.N., 2014. Texture-based fruit detection. Precision Agric., 1–22

Chinchuluun, R., Lee, W.S., Burks, T.F., 2007. Machine vision-based citrus yield mapping system. In: Proc. Fla. State Hort. Soc., (119), pp. 142–147.

Cohen, O., Linker, R., Naor, A., 2011. Estimation of the number of apples in color images recorded in orchards. Comput. Comput. Technol. Agric. IV, 630–642.

Dahikar, M.S.S., Rode, S.V., 2014. Agricultural crop yield prediction using artificial neural network approach. Int. J. Innovative Res. Electr. Electron. Instrum. Control Eng. 2 (1), 683–686.

D'Esnon, A.G., Rabatel, G., Pellenc, R., Journeau, A., Aldon, M.J., 1987. Magali: A Self-Propelled Robot to Pick Apples. ASAE Paper 87–1037, St. Joseph, MI.

Dobrusin, Y., Edan, Y., Grinshpun, J., Peiper, U.M., Hetzroni, A., 1992. Real-Time Image Processing for Robotic Melon Harvesting. ASAE Paper No. 92-3515, St. Joseph, MI.

Dobrusin, Y., Edan, Y., Grinshpun, J., Peiper, U.M., Wolf, I., Hetzroni, A., 1993. Computer image analysis to locate targets for an agricultural robot. Comput. Anal. Images Patterns, 775–779.

Duda, R.O., Hart, P.E., Stork, D.G., 2012. Pattern Classification. John Wiley & Sons.

Edan, Y., Rogozin, D., Flash, T., Miles, G.E., 2000. Robotic melon harvesting. IEEE Trans. Rob. Autom. 16 (6), 831–835.

Fathallah, F.A., 2010. Musculoskeletal disorders in labor-intensive agriculture. Appl. Ergon. 41 (6), 738–743.

Feng, J., Liu, G., Wang, S., Zeng, L., Ren, W., 2012. A novel 3D laser vision system for robotic apple harvesting. In: ASABE Annual International meeting presentation. Paper Number: 12-1341025, St. Joseph, MI.

Fennimore, S.A., Doohan, D.J., 2008. The challenges of specialty crop weed control, future directions. Weed Technol. 22 (2), 364–372.

Font, D., Pallejà, T., Tresanchez, M., Runcan, D., Moreno, J., Martínez, D., Palacín, J., 2014. A proposal for automatic fruit harvesting by combining a low cost stereovision camera and a robotic arm. Sensors 14 (7), 11557–11579.

Gallardo, R.K., Taylor, M.R., Hinman, H., 2010. Cost Estimates of Establishing and Producing Gala Apples in Washington. Extension Fact Sheet FS005E, University of Washington, School of Economic Sciences, Tree Fruit Research and Extension Center, Wenatchee, WA.

Gongal, A. 2014. Over-the -Row Machine Vision System for Improved Apple Crop-Load Estimation. MSc. Thesis (Chapter 3), Washington State University, USA.

Gongal, A., Amatya, S., Karkee, M. 2014. Identification of repetitive apples for improved crop-load estimation with dual-side imaging. In: ASABE Annual International Meeting, Paper Number: 141888882, St. Joseph, MI.

Gongal, A., Silwal, A., Amatya, S., Karkee, M., Zhang, Q., Lewis, K., 2015. Apple crop-load estimation with over-the-row machine vision system. Comput. Electron. Agric. (submitted for publication).

Gonzalez, C.R., Woods, R.E., Eddins, S.L. (Eds.), 2010. Digital Image Processing Using Matlab, second ed. McGraw Hill Companies, New Delhi.

Hannan, M.W., Burks, T.F., 2004. Current developments in automated citrus harvesting. In: ASAE/CSAE Annual International Meeting. Paper Number: 043827, St. Joseph, MI.

Hannan, M.W., Burks, T.F., 2007. A real-time machine vision algorithm for robotic citrus harvesting. In: ASABE Annual International Meeting. Paper Number: 073125. St. Joseph, MI.

Hannan, M.W., Burks, T.F., Bulanon, D.M., 2010. A machine vision algorithm combining adaptive segmentation and shape analysis for orange fruit detection. Agric. Eng. Int.: CIGR J. 1281 (11), 1–17 (Manuscript).

Harrell, R.C., Adsit, P.D., Munilla, R.D., Slaughter, D.C., 1990. Robotic picking of citrus. Robotica 8 (04), 269–278.

Harrell, R.C., Slaughter, D.C., Adsit, P.D., 1985. Vision guidance of a robotic tree fruit harvester. In: 1985 Cambridge Symposium. International Society for Optics and Photonics, pp. 537–545.

Harrell, R.C., Slaughter, D.C., Adsit, P.D., 1989. A fruit-tracking system for robotic harvesting. Mach. Vis. Appl. 2 (2), 69–80.

Hofmann, J., Snyder, K., Keifer, M., 2006. A descriptive study of workers' compensation claims in Washington State orchards. Occup. med. 56 (4), 251–257.

Huang, Y., Lan, Y., Thomson, S.J., Fang, A., Hoffmann, W.C., Lacey, R.E., 2010. Development of soft computing and applications in agricultural and biological engineering. Comput. Electron. Agric. 71 (2), 107–127.

Hunter, R.S., 1948. Photoelectric Color-Difference Meter. In: Proceedings of the Winter Meeting of the Optical Society of America. JOSA 38 (7), pp. 661.
Ji, W., Zhao, D., Cheng, F., Xu, B., Zhang, Y., Wang, J., 2012. Automatic recognition vision system guided for apple harvesting robot. Comput. Electr. Eng. 38, 1186–1195.
Jiménez, A.R., Ceres, R., Pons, J.L., 1999b. A machine vision system using a laser radar applied to robotic fruit harvesting. In: Computer Vision Beyond the Visible Spectrum: Methods and Applications, pp.110–119.
Jiménez, A.R., Jain, A.K., Ceres, R., Pons, J.L., 1999a. Automatic fruit recognition: a survey and new results using range/attenuation images. Pattern Recogn. 32 (10), 1719–1736.
Juste, F., Fornes, F., Plá, F., Molto, E., Blay, F., 1991. Primeros resultados en campo de un prototipo de brazo robotizado para la recolección de cıtricos. In: CIMA91, 23 Conf. Int. Maquinaria Agrıcola, Zaragoza, pp. 433–440.
Kane, K., Lee, W., 2006. Spectral sensing of different citrus varieties for precision agriculture. In: ASABE Annual International Meeting, Paper Number: 061065, St. Joseph, MI.
Kane, K.E., Lee, W.S., 2007. Multispectral imaging for in-field green citrus identification. In: ASABE Annual International Meeting, Paper Number: 073025. St. Joseph, MI.
Karkee, M., Adhikari, B., Amatya, S., Zhang, Q., 2014. Identification of pruning branches in tall spindle apple trees for automated pruning. Comput. Electron. Agric. 103, 127–135.
Karkee, M., Zhang, Q., 2012. Mechanization and automation technologies in specialty crop production. ASABE Resour. Mag., 16–17 (September/October)
Kim, Y., Reid, J., 2004. Apple yield mapping using a multispectral imaging sensor. In: International Conference on Agricultural Engineering, AgEng2004 Paper. Number: 010-PA- 235.
Kitamura, S., Oka, K., 2005. Recognition and cutting system of sweet pepper for picking robot in greenhouse horticulture. In: IEEE International Conference on Mechatronics and Automation. Niagara Falls, Canada.
Klimentjew, D., Hendrich, N., Zhang, J., 2010. Multi sensor fusion of camera and 3D laser range finder for object recognition. In: IEEE Conference Multisensor Fusion and Integration for Intelligent Systems (MFI), pp. 236–241.
Kondo, N., Kawamura, N., 1983. Method of detecting fruit by visual sensor attached to manipulator. Agric. Mach. 13, 45–52.
Kondo, N., Monta, M., Fujiura, T., 1996. Fruit harvesting robots in Japan. Adv. Space Res. 18 (1), 181–184.
Kurtulmus, F., Lee, W.S., Vardar, A., 2011. Green citrus detection using 'EigenFruit', color and circular Gabor texture features under natural outdoor conditions. Comput. Electron. Agric. 78 (2), 140–149.
Kurtulmus, F., Lee, W.S., Vardar, A., 2014. Immature peach detection in color images acquired in natural illumination conditions using statistical classifiers and neural network. Precision Agric. 15 (1), 57–79.
Labayrade, R., Royere, C., Gruyer, D., Aubert, D., 2005. Cooperative fusion for multi-obstacles detection with use of stereovision and laser scanner. Auton. Robots 19 (2), 117–140.
Li, P., Lee, S.H., Hsu, H.Y., 2011. Review on fruit harvesting method for potential use of automatic fruit harvesting systems. Proc. Eng. 23, 351–366.
Liang, D., Weng, K., Wang, C., Liang, G., Chen, H., Wu, X., 2014. A 3D object recognition and pose estimation system using deep learning method. In: 4th IEEE International Conference In Information Science and Technology (ICIST), pp. 401–404.
Linker, R., Cohen, O., Naor, A., 2012. Determination of the number of green apples in RGB images recorded in orchards. Comput. Electron. Agric. 81, 45–57.
Lindner, M., Kilb, A., Hartmann, K., 2007. Data-fusion of pmd-based distance information and high-resolution RGB-images. In: IEEE Sym. Signals Circuits & Systems (ISSCS), pp. 121–124.
Lowe, D.G., 1999. Object recognition from local scale-invariant features. In: Proceedings of the International Conference on Computer Vision, 2, pp. 1150–1157.
Lowe, D.G., 2004. Distinctive image features from scale invariant features. Int. J. Comput. Vision 60 (2), 91–110.
Mehta, S.S., Burks, T.F., 2014. Vision-based control of robotic manipulator for citrus harvesting. Comput. Electron. Agric. 102, 146–158.
Nakarmi, A.D., Tang, L., 2012. Automatic inter-plant spacing sensing at early growth stages using a 3D vision sensor. Comput. Electron. Agric. 82, 23–31.
Okamoto, H., Lee, W.S., 2009. Green citrus detection using hyperspectral imaging. Comput. Electron. Agric. 66 (2), 201–208.
Parish, E., Goksel, A.K., 1977. Pictorial pattern recognition applied to fruit harvesting. Trans. ASAE 20, 822–827.
Payne, A.B., Walsh, K.B., Subedi, P.P., Jarvis, D., 2013. Estimation of mango crop yield using image analysis–segmentation method. Comput. Electron. Agric. 91, 57–64.
Payne, A., Walsh, K., Subedi, P., Jarvis, D., 2014. Estimating mango crop yield using image analysis using fruit at 'stone hardening' stage and nighttime imaging. Comput. Electron. Agric. 100, 160–167.
Pejsa, J.H., Orrock, A.K., 1984. Intelligent robot systems; potential agricultural application. In: 1st International Conference on Robotics and Intelligent Machines in Agriculture, ASAE paper 4–84, St Joseph, MI 49085, pp. 104–111.
Peterson, D.L., Bennedsen, B.S., Anger, W.C., Wolford, S.D., 1999. A systems approach to robotic bulk harvesting of apples. Trans. ASAE–Am. Soc. Agric. Eng. 42 (4), 871–876.
Pla, F., Juste, F., Ferri, F., 1993. Feature extraction of spherical objects in image analysis: an application to robotic citrus harvesting. Comput. Electron. Agric. 8 (1), 57–72.
Plebe, A., Grasso, G., 2001. Localization of spherical fruitsfruit for robotic harvesting. Mach. Vis. Appl. 13 (2), 70–79.
Qiang, L., Jianrong, C., Bin, L., Lie, D., Yajing, Z., 2014. Identification of fruit and branch in natural scenes for citrus harvesting robot using machine vision and support vector machine. Int. J. Agric. Biol. Eng. 7 (2), 115–121.
Rakun, J., Stajnko, D., Zazula, D., 2011. Detecting fruits in natural scenes by using spatial-frequency based texture analysis and multiview geometry. Comput. Electron. Agric. 76 (1), 80–88.
Regunathan, M., Lee, W.S., 2005. Citrus fruit identification and size determination using machine vision and ultrasonic sensors. In: ASABE Annual International Meeting, Paper Number: 053017, St. Joseph, MI.
Rumelhart, D., Hinton, G., Williams, R., 1986. Learning internal representations.
Safren, O., Alchanatis, V., Ostrovsky, V., Levi, O., 2007. Detection of green apples in hyperspectral images of apple-tree foliage using machine vision. Trans. ASABE 50 (6), 2303–2313.
Sarig, Y., 1993. Robotics of fruit harvesting: a state-of-the-art review. J. Agric. Eng. Res. 54 (4), 265–280.
Scarfe, A.J., Flemmer, R.C., Bakker, H.H., Flemmer, C.L., 2009. Development of an autonomous kiwifruit picking robot. In: 4th International Conference on Autonomous Robots and Agents. Wellington, New Zealand.
Schertz, C.E., Brown, G.K., 1968. Basic considerations in mechanizing citrus harvest. Trans. ASAE, 343–346.
Seng, W.C., Mirisaee, S.H., 2009. A new method for fruits recognition system. Int. Conf. Electr. Eng. Inf. 1, 130–134.
Sengupta, S., Lee, W.S., 2012. Identification and determination of the number of green citrus fruit under different ambient light conditions. In: International Conference of Agricultural Engineering CIGR-AgEng2012.
Shapiro, L.G., Stockman, G.C., 2001. Computer Vision. Prentice Hall, Upper Saddle River, NJ.
Silwal, A., Gongal, A., Karkee, M., 2014. Apple identification in field environment with over-the- row machine vision system. Agric. Eng. Int.: CIGR J. 6 (4), 66–75.
Sistler, F., 1987. Robotics and intelligent machines in agriculture. IEEE J. Rob. Autom. 3 (1), 3–6.
Sites, D., 1988. Computer vision to locate fruit on tree. Trans. ASAE 31 (1), 257–263.
Slaughter, D.C., Harrell, R.C., 1989. Discriminating fruit for robotic harvest using color in natural outdoor scenes. Trans. ASAE 32 (2), 757–763.
Stajnko, D., Lakota, M., Hočevar, M., 2004. Estimation of number and diameter of apple fruits in an orchard during the growing season by thermal imaging. Comput. Electron. Agric. 42 (1), 31–42.
Stajnko, D., Rakun, J., Blanke, M., 2009. Modelling apple fruit yield using image analysis for fruit color, shape and texture. Eur. J. Horticult. Sci. 74 (6), 260–267.
Tabb, A.L., Peterson, D.L., Parker, J., 2006. Segmentation of Apple fruit from video via background modeling. An ASABE Meeting Presentation, Paper Number: 063060, St. Joseph, MI.
Tamura, H., Mori, S., Yamawaki, T., 1978. Texture features corresponding to visual perception. IEEE Trans. Syst. Man Cyber. SMC-8 (6), 460–473.
Tanigaki, K., Fujiura, T., Akase, A., Imagawa, J., 2008. Cherry-harvesting robot. Comput. Electron. Agric. 63, 65–72.
Tong, J., Zhang, Q., Karkee, M., Jiang, H., Zhou, J. 2014. Understanding the Dynamics of Hand Picking Patterns of Fresh Market Apples. ASABE Paper No.141898024. St. Joseph, Mich.
Van Henten, E.J., Hemming, J., Van Tuijl, B.A.J., Kornet, J.G., Meuleman, J., Bontsema, J., Van Os, E.A., 2002. An autonomous robot for harvesting cucumbers in greenhouses. Auton. Robots 13 (3), 241–258.
Wachs, J.P., Stern, H.I., Burks, T., Alchanatis, V., 2010. Low and high-level visual feature-based apple detection from multi-modal images. Precision Agric. 11 (6), 717–735.
Wachs, J., Stern, H.I., Burks, T., Alchanatis, V., Bet-Dagan, I., 2009. Apple detection in natural tree canopies from multimodal images. In: Proceeding of the Joint International Agricultural Conference, JIAC, pp. 293–302.
Wang, J.J., Zhao, D.A., Ji, W., Tu, J.J., Zhang, Y., 2009. Application of support vector machine to apple recognition using in apple harvesting robot. In: ICIA'09. International Conference In Information and Automation, pp. 1110–1115.
Wang, Q., Nuske, S., Bergerman, M., Singh, S., 2013. Automated crop yield estimation for apple orchards. Exp. Robotics, 745–758.
Whittaker, A.D., Miles, G.E., Mitchell, O.R., Gaultney, L.D., 1987. Fruit location in a partially occluded image. Trans. ASAE-Am. Soc. Agric. Eng. (USA).
Yamamoto, K., Guo, W., Yoshioka, Y., Ninomiya, S., 2014. On plant detection of intact tomato fruits using image analysis and machine learning methods. Sensors 14 (7), 12191–12206.
Yang, L., Dickinson, J., Wu, Q.M.J., Lang, S., 2007. A fruit recognition method for automatic harvesting. In: 14th International Conference in Mechatronics and Machine Vision in Practice, pp. 152–157.
Yu, J., Weng, K., Liang, G., Xie, G., 2013. A vision-based robotic grasping system using deep learning for 3D object recognition and pose estimation. IEEE Int. Conf. Robotics Biomimetics (ROBIO), 1175–1180.
Zhang, J., Tan, T., Ma, L., 2002. Invariant texture segmentation via circular Gabor filters. In: Proceedings of the 16th International Conference on Pattern Recognition. IEEE Computer Society, Quebec City, Quebec, Canada, (2), pp. 901–904.
Zhao, D.A., Jidong, L., Wei, J., Ying, Z., Yu, C., 2011. Design and control of an apple harvesting robot. Biosyst. Eng. 110 (2), 112–122.
Zhao, J., Tow, J., Katupitiya, J., 2005. On-tree fruit recognition using texture properties and color data. IEEE/RSJ Int. Conf. Intell. Robots Syst., 263–268
Zhou, R., Damerow, L., Sun, Y., Blanke, M.M., 2012. Using color features of cv. 'Gala' apple fruits in an orchard in image processing to predict yield. Precision Agric. 13 (5), 568–580.