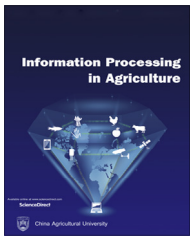
Available at www.sciencedirect.com

INFORMATION PROCESSING IN AGRICULTURE 5 (2018) 498–503

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



Apple fruit size estimation using a 3D machine vision system

A. Gongal, M. Karkee *, *S. Amatya*

Biological Systems Engineering Department, United States
Center for Precision and Automated Agricultural Systems, United States

ARTICLE INFO

Article history:
Received 22 December 2017
Received in revised form
29 May 2018
Accepted 5 June 2018

Keywords:
Machine vision
Fruit identification
3D registration
Size estimation
Major axis

ABSTRACT

Estimation of fruit size in tree fruit crops is essential for selective robotic harvesting and crop-load estimation. Machine vision systems for fruit detection and localization have been studied widely for robotic harvesting and crop-load estimation. However, only a few studies have been carried out to estimate fruit size in orchards using machine vision systems. This study was carried out to develop a machine vision system consisting of a color CCD camera and a time-of-flight (TOF) light-based 3D camera for estimating apple size in tree canopies. As a measure of fruit size, the major axis (longest axis) was estimated based on (i) the 3D coordinates of pixels on corresponding apple surfaces, and (ii) the 2D size of individual pixels within apple surfaces. In the 3D coordinates-based method, the distance between pairs of pixels within apple regions were calculated using 3D coordinates, and the maximum distance between all pixel pairs within an apple region was estimated to be the major axis. The accuracy of estimating the major axis using 3D coordinates was 69.1%. In the pixel-size-based method, the physical sizes of pixels were estimated using a calibration model developed based on pixel coordinates and the distance to pixels from the camera. The major axis length was then estimated by summing the size of individual pixels along the major axis of the fruit. The accuracy of size estimation increased to 84.8% when the pixel size-based method was used. The results showed the potential for estimating fruit size in outdoor environments using a 3D machine vision system.

1. Introduction

Precision agricultural technologies have helped producers manage their farms effectively, minimizing production costs by optimizing the use of land, water, seeds, fertilizers, pesticides, labor, and other farming inputs. This technology also improves crop productivity and the performance of farm machinery, while simultaneously reducing dependency on human labor.

Precision agricultural technologies such as variable rate fertilizer application [1], weed control [2–4] and crop monitoring [5] have widely been investigated for row crops. Yield estimation and mapping are both key to the successful application of these precision agricultural technologies to various crop production operations. However, when it comes to tree fruit crops such as apples and citrus, application of precision agricultural technologies and practices is limited. Among other factors, the lack of accurate crop-load estimation and yield mapping techniques explains why precision

* Corresponding author at: 24106 N Bunn Rd, Prosser, WA 99350, United States.
E-mail address: manoj.karkee@wsu.edu (M. Karkee).
Peer review under responsibility of China Agricultural University.
https://doi.org/10.1016/j.inpa.2018.06.002

agricultural technology and practices have not yet been applied to tree fruit crops.

Fruit counting and size estimation are two critical aspects of crop-load estimation and yield mapping in tree fruit crops. Several researchers have investigated fruit identification using machine vision systems for robotic harvesting and crop-load estimation [6–14]. Occlusion of fruits due to leaves, branches and other fruits, and variable lighting conditions are some of the prime factors that make it more challenging to achieve desired accuracy and robustness in fruit detection [6]. To minimize error due to variable lighting condition, different studies have been carried out in night time environments and while using tunnel structures around tree canopies. To avoid fruit occlusion, tree canopies are imaged from both front and back using multiple sensors [15–17], which helps to increase fruit detection accuracy. Although fruit size estimation is required for selective robotic harvesting of mature and good-sized fruits, only a few studies [18–21] have investigated fruit size estimation using a machine vision system. Wang et al. [18] used an RGB-D camera for size estimation of mango fruits in trees. The authors used thin lens theory to estimate fruit size. Regunathan and Lee [19] used three image classification techniques for fruit segmentation in color images along with range information from an ultrasonic sensor to estimate the size of citrus fruit. However, the size estimation method was not described and a root mean square error of 0.4 cm, with a standard deviation of 1.6 cm, was reported in this study. Stanjko and Čmelik [20] used thermal imaging to estimate apple size during the growing season. The authors found a correlation coefficient (R^2) of 0.78 between manual measurement and the estimated diameter of the apples. Accuracy was low during the early growing season in comparison to the harvesting period, as small or immature fruits radiated less heat.

In the study, we focused on estimating the size of apples in tree canopies trained to a modern canopy architecture using an over-the-row machine vision system consisting of a color camera and a 3D camera. The accuracy of size estimation methods depends on the accuracy of fruit segmentation and detection methods. Therefore, a tunnel structure was created under the over-the-row platform using farbric covers to avoid variable lighting conditions, which is one of the major sources of error in fruit identification reported earlier [6,7,22,23]. A color charged coupled device (CCD) camera was used to acquire apple canopy images; these images were then used to segment apples out and estimate their features. A 3D photonic mixing device (PMD) camera was used to acquire 3D information on the segmented apples. The specific objective of this study was to develop two different methods (3D coordinate-based and pixel-size-based methods) to estimate the size of apples in outdoor environments and compare their accuracies.

2. Materials and methods

2.1. Machine vision system

Color images of apple canopies were acquired using a CCD camera (Prosilica GigE 1290c, Allied Vision Technologies, Stadtroda, Germany). In addition, 3D images were acquired using a time-of-flight (TOF) light-based camera, also called a 3D camera (PMD CamCube 3.0, PMD Technologies, Siegen, Germany). Both cameras were installed together (Fig. 1a) on an over-the-row sensing platform to simultaneously capture canopy images with a similar field of view. To provide a controlled lighting environment, white LED lights with a color temperature of 5116 K (Daylight) (Fig. 1b) (Trilliant® 36 Light Emitting Diode Grote, Madison, Indiana) were installed and a tunnel structure (Fig. 1c) was created on the platform by covering and surrounding it with opaque materials (more details in [15]).

2.2. Field data collection

Field data were collected in commercial apple orchards trained to tall spindle architecture in Grandview, Washington (WA) with the cameras and platform discussed above in Section 2.1. Color and 3D images of "Fuji" apples (Fig. 2) were captured just before they were harvested. As the sizes of partially hidden apples in images will be different from the sizes obtained if the apples were completely visible, the size predictions of occluded apples will include larger errors. To avoid errors in size estimation caused by partially hidden apples, only completely visible apples (from at least one side of the canopy) were considered in this study, with an assumption that those apples will represent average fruit size and fruit size distribution in the trees. For this study, 150 apples were selected based on their visibility from image of 25 randomly selected trees. The major and minor axes of these apples were measured using a Vernier Caliper as a ground truth. In addition, those apples were labeled on the trees to ensure that the same apples were used in comparing actual and estimated fruit size.

2.3. Camera calibration and apple size estimation

A camera calibration method was used to determine the intrinsic and extrinsic parameter of the color and 3D cameras. For this task, 25 checkerboard images were acquired at different positions and angles with the color and 3D cameras. Then a camera calibration toolbox provided by MATLAB [24] was used to determine the intrinsic and extrinsic camera parameters. Co-registration of the color and the 3D cameras in a common coordinate system was then used to estimate intrinsic and extrinsic camera parameters. These methods are described in detail in [15].

Image processing was implemented in MATLAB software (R2012a, Math Works Inc., Natick, Massachusetts) to segment and detect apples in color images using color, shape and size features, which are described in detail in [15]. First, histograms of these images were equalized in Hue, Saturation and Intensity (HSI) color space, which amplifies the color differences between apples (foreground), and leaves and branches (background) in the color images. A Wiener filter [25] of neighborhood size 5×5 was applied to filter out noises. Images were then converted back to red, green and blue (RGB) color space, and then to grayscale images by subtracting the green and blue channels from the red channel. Otsu's thresholding method [25] was then applied to create binary images by separating image pixels into foreground and background classes. Otsu's method performs automated thresholding

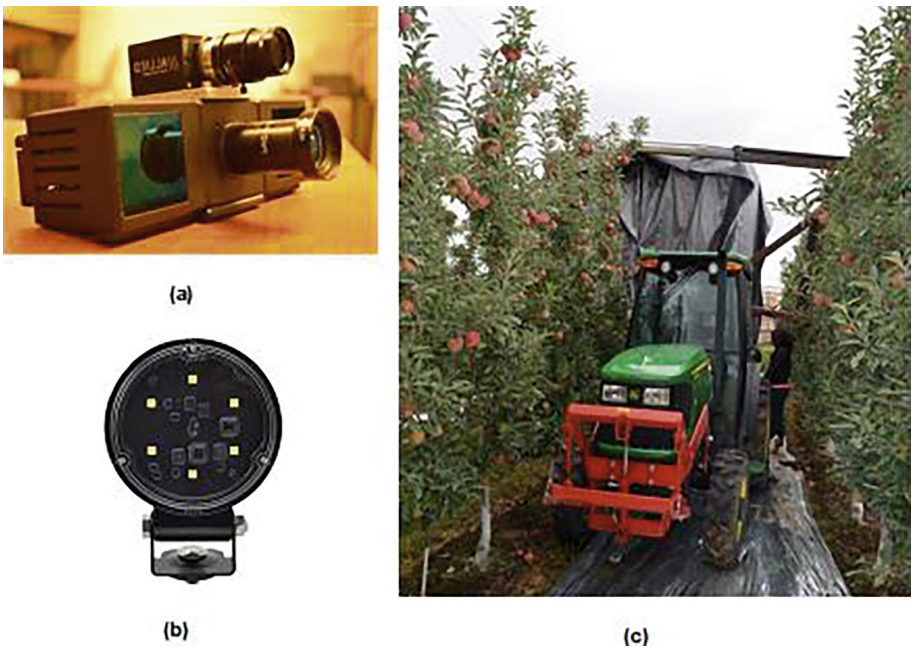


Fig. 1 – Cameras and platform used for the field data collection in this study; (a) A CCD color camera installed together with a 3D Camera; (b) LED light; and (c) Over-the-Row (OTR) sensor platform with a tunnel structure.

Fig. 2 – An image of apples captured in a commercial 'Fuji' apple orchard (Grandview, WA).

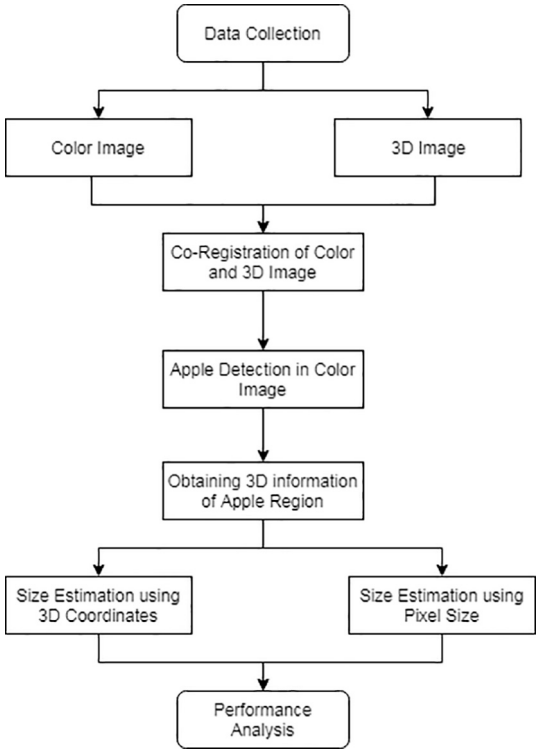


Fig. 3 – Functional block diagram of apple size estimation methods.

using the histogram of a grayscale image and the optimal gray level threshold to divide the images into forground and background pixels or objects. At the end of this process, the Circular Hough Transform (CHT) [25] was used to identify circular objects (apples) in these binary images.

The size of detected apples was estimated using two different methods (Fig. 3): (i) Apple sizing using 3D coordinates; and (ii) Apple sizing using the physical size of image pixels. These methods are described in detail in the following subsections.

2.3.1. *Apple sizing based on 3D coordinates*

Using co-registration of color and 3D images, three dimensional information for individual apple regions (in color images) was determined. This overlaying provided the 3D coordinates (x, y, and z) of each pixel within the apple regions. Then, the Euclidean distances between two individual pixels in an apple region were determined using 3D pixel coordinates. The maximum distance between any two pixels (out

of all possible pairs) within an apple region was calculated and used as the major axis of the corresponding apple. The estimated size of the major axis was then compared with the ground truth (measured manually using a Vernier Caliper) by calculating the mean absolute percentage error (MAPE) using equation (1).

$$MAPE = \frac{1}{N}\sum_{t=1}^{N}\frac{|L_a - L_e|}{L_a} \quad (1)$$

where

L_a = Actual length of major axis
L_e = Estimated length of major axis
N = Total number of apple samples

2.3.2. *Apple sizing based on pixel size*

Geometric properties of apple regions, such as area, major axis length and minor axis length, can be estimated easily in images in terms of the number of pixels. To covert these parameters into a physical unit, image resolution (the physical size of each pixel) was determined as discussed in the following paragraph. Then, the major axis length of each apple was estimated by adding the sizes of the individual pixels along the major axis of the fruit (Fig. 4).

To estimate the physical size or resolution of each pixel in an image, the relationship between pixel size, pixel coordinates and distance from the camera was modeled using a number of checkerboard images acquired while using both the color and 3D cameras. The experiment was conducted in the machine vision lab at the Center for Precision and Automated Agricultural Systems (CPAAS), Washington State University (WSU). A checkerboard (0.874 m × 0.782 m) was placed at ten different distances from the cameras, ranging from 0.6 m to 1.4 m. These distances for imaging were selected based on the practical ranges in the field tests that we conducted. Ten checker squares of a specific size (4.6 cm × 4.6 cm) were randomly selected from the checkerboard, and the distance to their centroids from the camera were measured using a laser sensor (DLR 130 K, Bosch, Stuttgart, Germany). Then, the actual pixel size within the checker square was determined using the physical size of the square and number of pixels within the square in the corresponding image. A regression model (Eq. (2)) was developed to predict pixel size based on pixel coordinates and distance to the center of the checker square from the camera. Fig. 5 is the plot of actual pixel size against predicted pixel size of checkerboard

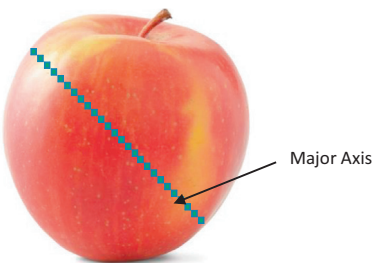


Fig. 4 – Estimation of major axis of apples as a sum of the physical resolution of pixels along the major axis of the fruit.

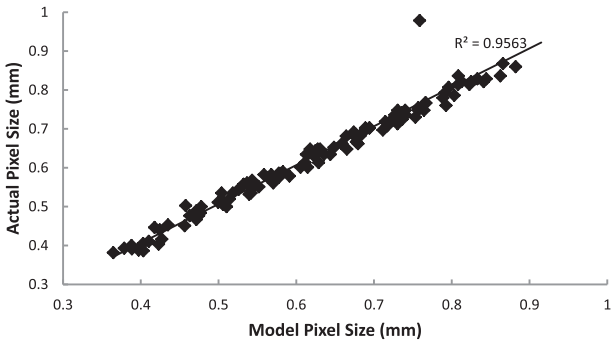


Fig. 5 – Actual pixel size vs modeled (estimated) pixel size of the checkerboard squares.

images using the model. The coefficient of regression (R^2) between actual pixel size and predicted pixel size was 0.96 and the mean absolute error was 2.3%.

$$s = ax + by + cd + ex^2 + fy^2 + gd^2 + h \quad (2)$$

where

s = pixel size;
x, y = pixel coordinates;
d = distance to pixel from camera; and
a, b, c, e, f, g, h = constants/co-efficient

For apple detection, image processing was done in color as described in Section 3. Co-registration of color and 3D images (described above) resulted in the measure of the distance of apple pixels from the camera and the coordinates of pixels along the major axis. Using the regression model (Eq. (2)), the sizes of all pixels along the major axis were estimated. The major axis of each apple was then estimated in physical units by summing the physical size of all the pixels along the major axis of apple region in question. The result was compared with the actual length (manually measured in the field) and the mean absolute percentage error (MAPE; Eq. (1)) was calculated. The accuracy of this method (in terms of MAPE) was compared with the accuracy obtained from apple sizing based on 3D coordinates method.

3. Results and discussions

The actual sizes and the estimated sizes of apples using the two different methods (Section 2.3) are shown in Table 1. The mean absolute error in estimating the major axis of apples based on 3D coordinates was 30.9% (Table 1). Fruit size, in general, was overestimated in comparison to actual fruit size using this method. The mean absolute error in estimating the major axis of apples based on pixel size was 15.2% (Table 1). Thus, apple size estimation accuracy when estimated based on the pixel size (84.8%) was substantially higher in comparison to estimates based on 3D coordinates (69.1%). The estimated fruit size with the pixel-size method was generally smaller in comparison to the actual size of the apples (Fig. 6).

The major sources of error in size estimation in both of the methods included errors involving apple segmentation, co-registration of 2D and 3D cameras and low resolution 3D

Table 1 – Comparison of accuracy of two different apple size estimation methods.

S.N	Actual Size (cm)	Estimated size (cm)		Absolute percentage error (%)	
		Using 3D coordinates	Using pixel size	Using 3D coordinates	Using pixel size
1	7.89	10.10	7.06	28.0	10.6
2	8.03	10.63	7.63	32.4	4.9
3	7.86	9.74	5.70	23.9	27.5
4	7.25	9.03	5.82	24.6	19.7
5	8.33	8.86	5.81	6.4	30.2
6	7.67	10.57	6.82	37.8	11.1
7	7.53	9.34	6.74	24.0	10.6
8	6.97	9.19	6.83	31.9	2.0
9	7.71	9.20	6.89	19.3	10.7
10	7.57	9.61	6.41	26.9	15.4
11	7.68	8.96	7.20	16.7	6.3
12	8.03	10.46	5.27	30.2	34.3
13	6.72	8.06	5.49	19.9	18.4
14	7.58	8.11	5.24	7.0	30.8
15	7.86	9.04	6.77	15.0	13.8
–	–	–	–	–	–
141	7.09	8.61	6.78	21.5	4.4
142	7.11	8.84	5.95	24.3	16.3
143	7.61	8.42	6.49	10.6	14.8
144	6.99	9.31	6.64	33.2	4.9
145	6.41	8.66	5.88	35.1	8.2
146	7.73	8.59	7.63	11.1	1.5
147	6.84	9.41	6.65	37.6	2.8
148	6.42	8.09	6.40	26.0	0.3
149	6.51	10.06	5.59	54.5	14.2
150	7.64	9.84	7.39	28.8	3.3
Mean absolute percentage error (%)				30.9%	15.2%

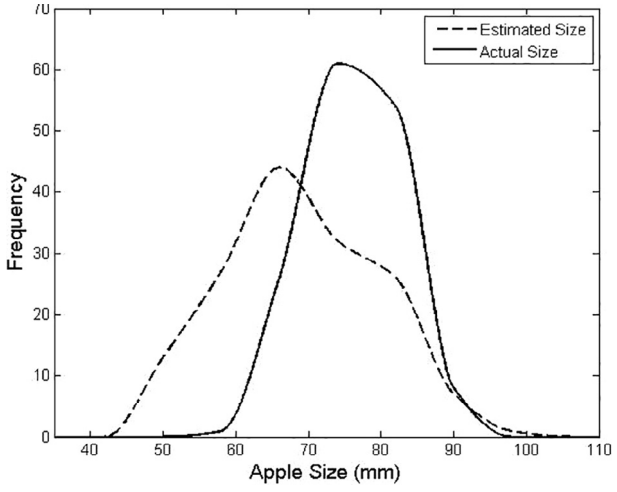


Fig. 6 – Distribution of actual size and predicted size of apples estimated using the pixel size in images.

Fig. 7 – Apples identified in a color image based on color, shape, and size features.

images. Errors of apple segmentation, primarily was from detecting only partial apple regions or mistaking the surroundings/background as apple regions (Fig. 7) which led to the inaccuracy on identifying apple pixels within the images. The difficulties inherent in overlaying 3D camera images on 2D color images resulted in errors in the 3D coordinates of apple regions. Further, the high level of inaccuracy in the apple size estimation based on 3D coordinates was due to the low resolution of the 3D camera images in comparison to the color camera images. Consequently, very little 3D coordinate information was available within individual apple regions. Thus, size estimation based on 3D coordinates would be erroneous if 3D information is not available for the pixels in apple boundaries. These errors in size estimation could potentially be reduced by improving the registration accuracy

between the 3D and color images. Also, improving image segmentation methods (e.g., using deep learning methods) and use of high resolution 3D cameras could have help improve size estimation accuracy.

4. Conclusions

A machine vision system to estimate the sizes of apples in tree canopies was developed. An over-the-row platform with a tunnel structure was used for imaging under controlled lighting conditions in order to improve apple segmentation accuracy. A color camera was used to capture images of apple tree canopies for apple detection and a 3D PMD camera was used to obtain 3D location information for apples. The major axes of the apples were then estimated using: (i) *3D coordinates of apple regions; and* (ii) *pixel size within apple regions*. The accuracy of estimating the major axis of apples was 69.1% for the first approach, based on the 3D coordinates of pixels in apple regions, and 84.8% for the second approach, based on pixel size.

Size estimation errors were mainly due to underlying errors in image segmentation and 3D mapping techniques, along with the low-resolution of the 3D sensor. Future research should address the need to improve size estimation with enhanced image processing methods for fruit segmentation (e.g., deep learning with improved 3D sensing and mapping techniques).

Conflict of interest

The authors declared that there is no conflict of interest.

Acknowledgements

This research was supported in part by the USDA's Hatch and Multistate Project Funds (Accession Nos. 1005756 and 1001246), the Washington Department of Agriculture, the Washington Red Raspberry Commission, and Washington State University's Agricultural Research Center (ARC). Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the U.S. Department of Agriculture or Washington State University.

REFERENCES

[1] Fleming KL, Westfall DG, Wiens DW, Brodahl MC. Evaluating farmer defined management zone maps for variable rate fertilizer application. Precis Agric 2000;2(2):201–15.
[2] Gerhards R, Oebel H. Practical experiences with a system for site-specific weed control in arable crops using real-time image analysis and GPS-controlled patch spraying. Weed Res 2006;46(3):185–93.
[3] Lee WS, Slaughter DC, Giles DK. Robotic weed control system for tomatoes. Precis Agric 1999;1(1):95–113.
[4] Slaughter DC, Giles DK, Downey D. Autonomous robotic weed control systems: a review. Comput Electron Agric 2008;61 (1):63–78.
[5] Chen X, Fang J, Gao W, HuangFu G, Lin K, Hu H. The research and application of agricultural planting monitoring system based on networking technology. Future Inform Eng 2014;49 (2):313–21.
[6] Cohen O, Linker R, Naor A. Estimation of the number of apples in color images recorded in orchards. Int Fed Inform Process AICT 2011;344:630–42.
[7] Bulanon D, Katoka T, Ota Y, Hiroma T. A color model of recognition of Apples by a Robotic harvesting system. J Jpn Soc Agric Mach 2002;64(5):123–33.
[8] Bulanon D, Katoka T. Fruit detection system and an end effector for robotic harvesting of Fuji apples. Agric Eng Int CIGR J 2010;12(1):203–10.
[9] Amatya S, Karkee M, Gongal A, Zhang Q, Whiting MD. Detection of cherry tree branches with full foliage in planar architecture for automated sweet-cherry harvesting. Biosyst Eng 2016;146:3–15.
[10] Ji W, Zhao D, Cheng F, Xu B, Zhang Y, Wang J. Automatic recognition vision system guided for apple harvesting robot. Comput Electr Eng 2012;38:1186–95.
[11] Linker R, Cohen O, Naor A. Determination of the number of green apples in RGB images recorded in orchards. Comput Electron Agric 2012;81:45–57.
[12] Safren O, Alchanatis V, Ostrovsky V, Levi O. Detection of green apples in hyperspectral images of apple tree foliage using machine vision. Trans ASABE 2007;50(6):2303–13.
[13] Linker R. A procedure for estimating the number of green mature apples in night-time orchard images using light distribution and its application to yield estimation. Precis Agric 2017;18:59–75.
[14] Qureshi WS, Payne A, Walsh KB, Linker R, Cohen O, Dailey MN. Machine vision for counting fruit on mango tree canopies. Precis Agric 2017;18(2):224–44.
[15] Gongal A, Silwal A, Amatya S, Karkee M, Zhang Q, Lewis K. Apple crop-load estimation with over-the-row machine vision system. Comput Electron Agric 2016;120:26–35.
[16] Silwal A, Gongal A, Karkee M. Apple identification in field environment with over the row machine vision system. Agric Eng Int CIGR J 2014;16(4):66–75.
[17] Wang Q, Nuske S, Bergerman M, Singh S. Automated crop yield estimation for apple orchards. Experim Robot 2013:745–58.
[18] Wang Z, Walsh KB, Verma B. On-tree mango fruit size estimation using RGB-D images. Sensors 2017;17(12):2738.
[19] Regunathan M, Lee WS. Citrus fruit identification and size determination using machine vision and ultrasonic sensors. In: ASABE annual international meeting, paper number: 053017, St Joseph, Michigan; 2005.
[20] Stajnko D, Čmelik Z. Modelling of apple fruit growth by application of image analysis. Agric Conspec Sci 2005;70 (2):59–64.
[21] Stajnko D, Lakota M, Hocevar M. Estimation of numbers and diameter of apple fruits in an Orchard during the growing season by thermal imaging. Comput Electron Agric 2004;42:31–42.
[22] Kurtulmus F, Lee WS, Vardar A. Green citrus detection using 'EigenFruit', color and circular Gabor texture features under natural outdoor conditions. Comput Electron Agric 2011;78 (2):140–9.
[23] Zhao J, Tow J, Katupitiya J. On-tree fruit recognition using texture properties and color data. In: IEEE/RSJ international conference on intelligent robots and systems; 2005. p. 263–8.
[24] Bouguet JY. Camera Calibration toolbox for MATLAB. Link: <http://www.vision.caltech.edu/bouguetj/calib_doc>; 2013 [2014].
[25] Gonzalez CR, Woods RE, Eddins SL. Digital image processing using matlab. 2nd ed. New Delhi: McGraw Hill Companies; 2010.