Precision Agric
DOI 10.1007/s11119-016-9443-z

Immature green citrus fruit detection and counting based on fast normalized cross correlation (FNCC) using natural outdoor colour images

Han Li[1,3] · **Won Suk Lee**[2] · **Ku Wang**[3]

Abstract A fast normalized cross correlation (FNCC) based machine vision algorithm was proposed in this study to develop a method for detecting and counting immature green citrus fruit using outdoor colour images toward the development of an early yield mapping system. As a template matching method, FNCC was used to detect potential fruit areas in the image, which was the very basis for subsequent false positive removal. Multiple features, including colour, shape and texture features, were combined in this algorithm to remove false positives. Circular Hough transform (CHT) was used to detect circles from images after background removal based on colour components. After building disks centred in centroids resulted from both FNCC and CHT, the detection results were merged based on the size and Euclidian distance of the intersection areas of the disks from these two methods. Finally, the number of fruit was determined after false positive removal using texture features. For a validation dataset of 59 images, 84.4 % of the fruits were successfully detected, which indicated the potential of the proposed method toward the development of an early yield mapping system.

Keywords Circular Hough transform · Colour component · Normalized cross correlation · Texture feature · Yield mapping

✉ Won Suk Lee
wslee@ufl.edu

[1] Beijing Research Center of Intelligent Equipment for Agriculture, Beijing Academy of Agriculture and Forestry Sciences, Beijing 100097, China

[2] Department of Agricultural & Biological Engineering, University of Florida, Gainesville, FL 32611, USA

[3] China Agricultural University, Beijing 100083, China

Published online: 12 March 2016

Introduction

Estimating early yield of immature green citrus fruit is very important for citrus growers. This information can help growers with preparation of packing and storage facilities, and estimation of harvesting logistics and expected cost. It can also help the growers to identify spatial yield variability of their citrus groves, which would further help them optimize their site-specific management and increase their profits. Florida produced about 70 % citrus of the US during the last decade (USDA-NASS 2011). As the production of citrus increased, greater accuracy of yield mapping would be desired. Until now, yield mapping is based on estimates and projections from manual counting and measurements of selected trees, which are simple in concept but complex in planning and management. The counting and measurements are also time consuming and labour intensive. There are four basic parameters in citrus production forecast: number of bearing trees, number of fruit per tree, fruit size, and fruit loss from droppage (USDA-NASS 2011). Among these four parameters, the first two have greatest impact on the forecast, and the number of fruit per tree is the key for production forecast. Therefore, to help the farmers get early yield mapping data, this research focused on studying how to determine the number of immature citrus fruit per tree based on machine vision toward the development of an early yield estimation system. A machine vision based system is an excellent choice because of its efficiency and capability of recognizing size, shape, colour, texture and numerical attributes of the objects (Shin et al. 2012). In comparing with the current method of manual fruit counting, the machine vision based system could help the farmers investigate larger citrus groves faster and more efficiently. This study focused on automatic immature green citrus fruit counting in natural outdoor red–green–blue (RGB) images.

Many studies have been reported to automatically detect immature green fruit in images. These studies usually fall into two categories based on the imaging systems used. One category is those studies using infrared (IR), multispectral, and hyperspectral imaging. Some of them described applications of thermal imaging for detecting objects or plant parts in a low contrast background because thermal imaging utilizes different thermal characteristics of observed objects compared with the use of visual cameras. For IR images, Sapina (2001) computed six texture features based on co-occurrence matrix, which is the second order statistics method characterizing spatial interrelationships of grey tones. The texture images showed good results of variance and correlation for all tested images, which were helpful to identify warm objects and background in low contrast images. Stajnko et al. (2004) developed an algorithm to estimate number and diameter of apple fruits in an orchard by thermal imaging during a growing season. The coefficients of determination, R2, were used to show the relationship between the manual measurement and estimated results. They reported that the R2 between the manually counted number of fruit and the number estimated was ranged from 0.83 to 0.88, and the R2 between the manually measured fruit'' diameter and the estimated diameter was ranged from 0.68 to 0.70 based on their algorithm. Some researchers have explored the feasibility of detecting fruits utilizing multispectral or hyperspectral information. Annamalai and Lee (2004) conducted spectral analysis for immature green citrus fruits and leaves using a spectrophotometer in a laboratory. They found two important wavelengths, 815 and 1190 nm, which were significant for fruit identification. Kane and Lee (2007) used a monochrome near-infrared camera with interchangeable optical band pass filters to capture images of green citrus fruit in Florida. The images were processed using indices and morphological operations. The results showed an R2 of 0.74 between the predicted number of citrus pixels and number of pixels

manually masked. Okamoto and Lee (2009) developed a hyperspectral image processing method to detect green citrus fruit in individual trees from hyperspectral images captured in 369–1042 nm. The detection success rates were 70–85 %, depending on the citrus varieties.

For immature green fruit detection, standard colour cameras have also been used. Wachs et al. (2010) developed a machine vision system aiming at recognizing occluded green apples within a tree canopy using both IR and colour images. Their apple detection algorithm was based on low-level and high-level visual features. Low-level features, which meant Haar wavelet features in the paper, were obtained in colour and thermal infrared images through an Adaboost algorithm. High-level features were colour and geometric properties obtained through image pre-processing, clustering and morphology operations. The results revealed that the low-level feature approach was superior to the high-level one, because it relied on shape descriptors rather than colour descriptors, which made it more robust to unconstrained illumination in natural scenes. And the recognition accuracy increased when the two approaches were combined. Zhou et al. (2012) developed apple recognition algorithms for two fruit development periods, i.e., after June drop and during ripening. A colour feature based method was used to estimate the number of apple fruits in an image after June drop, when the fruits were mostly young and small green fruitlets. This method yielded an R2 of 0.80 between the fruit counting algorithm and manually counted number of fruit. Linker et al. (2012) developed and validated an algorithm for estimating the number of apples in natural outdoor colour images. The four step method detected more than 85 % of the apples in the first dataset, while it detected close to 95 % of the actual apples in the second dataset of images, which were manually underexposed and recorded. Kurtulmus et al. (2011) developed a green citrus detection method using 'eigenfruit', colour and circular Gabor texture features under natural outdoor conditions. Gabor texture is a powerful tool in texture analysis, and the circular Gabor texture is a modified version of traditional Gabor texture, a circular symmetric version. They concluded that circular Gabor texture was a useful method for identifying fruit texture regions in the natural outdoor canopy. For a validation dataset, 75.3 % of the actual fruits were successfully detected using the proposed method. Then Kurtulmus et al. (2014) further developed an immature peach fruit detection method in colour images, which were also captured under natural outdoor conditions. Three methods were used to separate fruit objects in the images, including scanning an entire image with a sub-window process to extract the ratio of fruit object, extracting potential fruit regions with a morphological opening operation of binary image, and extracting potential fruit regions with a radial symmetry transform. Statistical classifiers, a neural network and a support vector machine classifier were built and used for detecting peach fruit, besides 'eigenfruit' and circular Gabor texture features. The proposed methods yielded correct detection accuracies of 71–85 % for a validation dataset. Bansal et al. (2013) proposed a fast Fourier transform (FFT) leakage based green citrus detection method, which yielded a correct detection rate of 82.2 %. A consumer grade digital colour camera was chosen to obtain images in this study because it is easier to operate and more affordable for growers, in comparison to multispectral or hyperspectral cameras. Also Sengupta and Lee (2014) conducted immature green citrus fruit identification and counting using similar images captured by an ordinary standard digital camera in a canopy under different light conditions. They developed a machine vision based algorithm which included two integral parts: shape analysis and texture classification. The green citrus detection accuracy they reported in a validation set of images was 80.4 %. Some researchers tried to use images captured by a mobile phone camera to detect mature fruits having different colour from the background.

For example, Gong et al. (2013) developed a new method based on an 8-connectedness chain code to estimate number of mature fruit, using images taken by an Android mobile phone. The performance of the fruit-counting algorithm achieved a recognition rate of 90 % for 40 validation images of citrus trees.

Compared with green fruit detection using infrared (IR), multispectral, and hyperspectral imaging, or non-green fruit detection using colour images, the task to be achieved in this study was much more difficult. It mainly has three challenges: colour similarity between fruits and leaves, non-uniform illumination, and partial occlusion. Some researchers have tried to solve these problems. Colour, texture, and shape were the most commonly used features for target classification or detection in machine vision. For example, Kurtulmus et al. (2011) used eigenfruit approach using intensity and saturation components, and circular Gabor texture. Pourreza et al. (2012) identified nine Iranian wheat seed varieties utilizing 131 textural features extracted. The average accuracy of 98.5 % was obtained. In this study, all of these mentioned features, colour, texture and shape, were used in the proposed method, to improve green fruit detection accuracy.

The objective for this study was to develop methods that can detect immature green citrus fruits accurately in the RGB images acquired outdoors by a standard digital colour camera. Figure 1 shows an example image captured by the camera, which was employed to develop the proposed algorithm. Specific objectives were:

1. to develop a method for removing background as much as possible based on colour component analysis,
2. to develop a fast normalized cross correlation (FNCC) based method to detect potential fruit positions, and
3. to identify as many green citrus as possible through combining colour, texture, and shape feature analysis.

Materials and methods

Image acquisition

Under natural daylight conditions, 118 digital colour images were acquired in October, 2010 from an experimental citrus grove in the University of Florida, Gainesville, Florida,

Fig. 1 An example RGB image containing immature green citrus fruit (Color figure online)

USA, using a consumer grade regular digital camera (PowerShot SD880IS, Canon Inc., USA). The citrus variety was Orlando Tangelo. The images were taken between 11:00 am and 4:00 pm, in 1 week period. The interval of the picture obtaining time was random. The weather conditions were sunny or cloudy, when the pictures were taken. The images were taken for both sunny and shadow side of the citrus canopy under natural daylight. The images were saved as 24-bit colour JPG format images with 3648 × 2736 pixels, corresponding to approximately 24 × 18 cm^2 actual scene. Matlab (Ver. R2010a, Mathworks, Natick, MA, USA) was used to realize the proposed green citrus detection algorithm in this study. Among the 118 images, 59 images were randomly selected as training images, and the other 59 were used as validation images. All the RGB images were resized to 912 × 684 pixels, using bilinear interpolation method, to improve computational efficiency.

Algorithm description

The flow charts of the proposed method are shown in Figs. 2 and 3. Figure 2 shows general steps for the method, which was divided into five steps. Firstly, FNCC (Lewis 1995) was applied to obtain potential fruit positions by calculating cross correlation between an input image and a template fruit image. A square was placed around the potential fruit position in the original RGB image based on the FNCC. The square obtained was called a "patch" in the rest of this study, and these specific patches were used as the study targets. Then false positive positions were filtered out through colour analysis for each patch. Thirdly, colour and shape feature were combined for detecting potential fruit positions utilizing circular Hough transform (CHT) (Duda and Hart 1972). R, B and H colour components were chosen to remove background (non-fruit objects), using the same colour component analysis method developed in Li et al. (2014), which will not be detailed in this paper. R, B was red and blue component in RGB image, respectively. H was hue component in HSV image, which was the result image after rgb2hsv colour space transformation was performed on the RGB image. Then the detection results from steps 2 and 3 were merged.

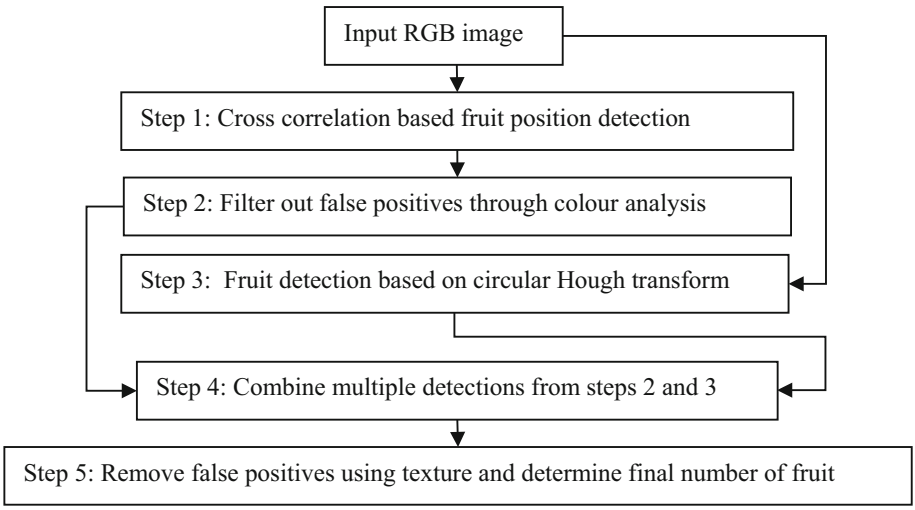

Fig. 2 General flow chart for the proposed algorithm

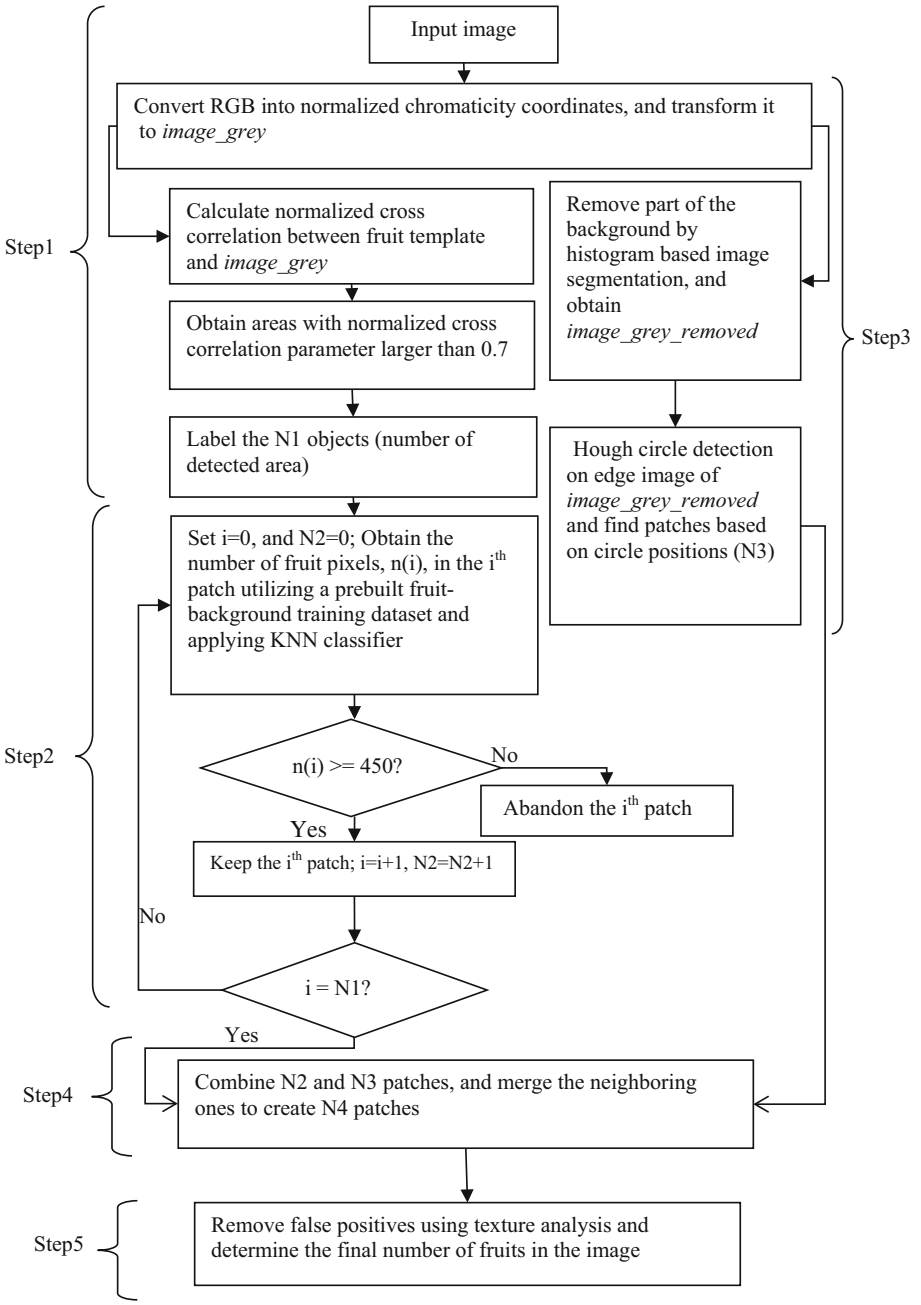

Fig. 3 Flow chart for the proposed algorithm with more details

Lastly, false positives were further filtered out through texture analysis for the remaining patches. Figure 3 shows more specific flow chart of the proposed method. In this figure, N1, N2, N3 and N4 indicate the total number of patches resulting from different steps. More specific details are discussed in the following subsections.

Normalized chromaticity coordinates

Conclusions could not be made directly based on non-normalized RGB pixel values of an object, because the non-normalized pixel values are highly sensitive to the illumination difference of the lighting source (Gonzalez and Woods 2002). Therefore, RGB images used in this study were transformed to normalized chromaticity coordinates first, since non-normalized RGB pixel values are proportional to total light reflected from a surface.

Normalized RGB coordinates could be obtained using the following equation (Woebbecke et al. 1995):

$$R = \frac{R_C}{R_M}, \quad G = \frac{G_C}{G_M}, \quad B = \frac{B_C}{B_M} \tag{1}$$

where R_C, G_C, B_C are the non-normalized values, and R_M, G_M, B_M are the maximum non-normalized RGB values.

The chromaticity coordinates is defined as the ratio of each of a set of three tristimulus values to their sum by CIE (Brill 2014). Hence, the chromatic coordinates for RGB colour space is defined by Eq. 2.

$$r = \frac{R}{R+G+B}, \quad g = \frac{G}{R+G+B}, \quad b = \frac{B}{R+G+B} \tag{2}$$

Since RM = GM = BM = 255 for 24-bit colour images, Eqs. 3 and 4 could be given through combing Eqs. 1 and 2:

$$r = \frac{R_C}{R_C+G_C+B_C}, \quad g = \frac{G_C}{R_C+G_C+B_C}, \quad b = \frac{B_C}{R_C+G_C+B_C} \tag{3}$$

$$r + g + b = 1 \tag{4}$$

After the images were normalized to chromaticity coordinates, the effect of illumination was reduced, because the normalized chromatic coordinates for shaded and unshaded surfaces were not as variable as non-normalized RGB values (Woebbecke et al. 1995).

FNCC based potential fruit position detection

After an RGB image was normalized to chromaticity coordinates, it was converted to grey scale image, which was named *image_grey*.

Fig. 4 Template image used for FNCC

FNCC is a template matching method for feature detection. It requires a typical template for the object to be detected. To obtain potential fruit positions in the image, FNCC coefficients were calculated between a fruit template and *image_grey*. The fruit template used in this study as shown in Fig. 4 was obtained by calculating an average for 20 small 30 × 30 pixel patches, which were randomly chosen from the training images.

This method normalizes the image and feature vectors to unit length, yielding a cosine similarity function like a correlation coefficient (Lewis 1995) shown as Eq. 5,

$$\gamma(u,v) = \frac{\sum_{x,y}\left[f(x,y) - \bar{f}_{u,v}\right]\left[t(x-u,y-v) - \bar{t}\right]}{\left\{\sum_{x,y}\left[f(x,y) - \bar{f}_{u,v}\right]^2 \sum_{x,y}\left[t(x-u,y-v) - \bar{t}\right]^2\right\}^{0.5}} \tag{5}$$

where f is an input grey scale image, and the sum in the numerator and denominator is over under the window containing the feature (the fruit template) positioned at u, v; γ is the correlation coefficient between the fruit template and *image_grey*; $\bar{t}$ is a mean of the template and $\bar{f}_{u,v}$ is a mean of $f(x, y)$ in the region under the template area. This equation is referred to as a normalized cross correlation.

The FNCC, which was proposed by Lewis (1995), has improved the computation efficiency greatly. It utilized the technique of computing a definite sum from a pre-computed running sum. And the algorithm could be found in Matlab as a function 'normxcorr2'. The resulting matrix contains correlation coefficients and its values may range from −1.0 to 1.0. For the training image dataset, nine cross correlation coefficient thresholds from 0.1 to 0.9 with a 0.1 interval were tried to determine the best threshold for green fruit detection. The best result was given when "0.7" was chosen as the threshold. Therefore, pixels with a coefficient higher than 0.7 were identified as potential fruit pixels. The total number of areas in the result image with a size bigger than 40 pixels was counted, and was assigned to N1 in Fig. 3. The threshold of 40 pixels was determined based on the trade-off analysis between the fruit detection accuracy and the algorithm running time, which indicated most of the areas smaller than 40 pixels were non-fruit objects.

Filter out false positives through colour analysis

For each of those N1 positions, a 30 × 30 pixel square patch located at the centroid of each area was cropped, and was used for colour analysis. The more pixels in each patch belong to a fruit class, the more likely that this patch belongs to a fruit. To identify how many pixels were fruit in each patch based on the colour information, fruit and background pixels needed to be investigated.

Therefore, a fruit-background pixel dataset including 10,000 fruit pixels, and 20,000 background pixels (10,000 leaf pixels and 10,000 all other background pixels) was built. This prebuilt data library was used for colour component selection. These pixels were randomly chosen from the 59 training images. A half of these pixels (5,000 + 10,000 pixels) were used for training and the other half were used for validation. The classification based on different colour components from different colour spaces can yield different classification results (Gonzalez and Woods 2002). Four commonly used colour spaces include red (® green (G), and blue (B); hue (H), saturation (S), and intensity (I); luminance (Y), blue-difference chroma (Cb), and red-difference chroma (Cr); and luminance (Y), in-phase (I), and quadrature-phase (Q). The training dataset was used to choose the best colour components, and the validation dataset was used to validate the efficiency of the chosen components. To choose which colour component subset could yield the best result,

a nonparametric measurement selection method, forward feature selection algorithm (FFSA) as described by Whitney (1971) and Kumar et al. (2001) was chosen. The general idea for this method is as following. Firstly the best single component with the highest accuracy is selected first, and then the best pair of features is selected where the pair includes the best single selected feature. A new component is added to the previously selected component(s) so that the classification accuracy increases maximally when it is added. This procedure is repeated with the rest of the features until the classification accuracy does not improve considerably or all components are used up. From the preliminary analysis by FFSA, the R, B, and H components were proved to be most efficient.

These three components, R, B, and H, of the training dataset were used in the proposed algorithm. Firstly, they were used for calculation of the number of fruit pixels in the N1 patches in step 1. Here N1 in Fig. 3 indicates the number of patches using FNCC. Then the histograms of the R, B, and H components of the training dataset were used in step 3 to help with the background removal. Using these three colour components, a K-nearest neighbour (KNN) (Duda et al. 2001) classifier was built and used for the pixel classification in step 2, because it performed best in the preliminary tests. If the pixel number in each patch was bigger than 450, which was a half of the patch size, this patch would be identified as a fruit, and it would be kept for further analysis. Otherwise, it was identified as background, and was abandoned. N2 in Fig. 3 indicates the total number of patches remained after colour analysis.

Colour based background removal and CHT based fruit detection

Since the colours of green leaves and green fruit are very similar, only colour information is not enough to identify fruit from the image (Kurtulmus et al. 2011; Bansal et al. 2013; Kurtulmus et al. 2014). Therefore, shape feature was also considered. Step 3 in Fig. 3 shows the general idea for shape based fruit detection. If grey images were directly used for circle detection, there would be many false positives. Thus, part of the background was removed based on histogram analysis of R, B, and H components of the image. The edge image of *image_grey_removed*, which was created in step 3, was obtained by edge detection using the 'canny' operator. Then the CHT (Duda and Hart 1972) was applied on the edge image of *image_grey_removed*, and N3 patches were found based on the centroids of circle positions.

The circular Hough transform was extended from the Hough transformation for line detection. A circle in the edge image could be described by Eq. 6. Then an arbitrary image point will be transformed into a surface in a parameter space defined by Eq. 7 (Duda and Hart 1972).

$$(x - a1)^2 + (y - b1)^2 = R1^2 \tag{6}$$

$$(x_i - a1)^2 + (y_i - b1)^2 = R1^2 \tag{7}$$

Each point in the edge image will be transformed into a right circular cone in a three-dimensional parameter space. A right circular cone is a cone whose axis is a line segment joining the vertex to the midpoint of the circular base. If the cones correspond to many figure points intersected at a single point, then all the figure points lie on the circle defined by those three parameters, which are a1, b1 and R1. Based on the size of fruits on the images, the minimum and maximum radius was set to be 60 and 180 pixels, respectively.

The maximum number of circles found in an image was set to be 15 because it was the maximum number of fruits found in the captured images.

Merging multiple detections

N2 and N3 potential fruit positions were found in steps 2 and 3, respectively, as shown in Fig. 3. Multiple detections might occur for a single fruit, which will introduce false positives. Therefore, merging was necessary for multiple detections.

For merging, a very simple yet efficient procedure was used. Firstly, three binary images initialized with the same size of the input image (912 $\times$ 684 pixels) were built. The value of all the pixels on them was initialized as '0'. Then, disks with a radius of 60 pixels were built corresponding to N2 and N3 centroids in the first two binary images, respectively, and the value of the pixels on the disks was set to '1'. The threshold of 60 pixels was chosen because it was approximately the minimum fruit radius found in the images, as mentioned in the last section. It means that if two fruits have more than one half overlapped, they will be identified as one since the distance between the overlapped fruit is smaller than 60 pixels. Then the intersection for these two images was calculated in the third binary image. The areas in the third binary image bigger than 600 pixels were identified as fruit. Here, "600" was a parameter determined based on trial and error from the training set images.

Further false positive removal and determination of the final number of fruit

Since there were still some false positives occurring in the result after step 4, some texture features were utilized for further false positive removal. Two texture features, smoothness (S) and entropy (E), were calculated for each patch corresponding to the N4 centroids obtained from step 4. Smoothness was chosen because most of the fruits had smoother surface than leaves and other background. And entropy was chosen based on a preliminary test, which indicated the patches on the fruits had lower entropy value compared with other patches with multiple leaves or twigs in them. The features were extracted from normalized histogram of the matrix for each patch in *image_grey*, which were calculated through Eqs. 8 and 9 (Pourreza et al. 2012), respectively.

$$S = 1 - \frac{1}{1 + \sigma^2} \tag{8}$$

where $\sigma = \sqrt{\sum_i i - \mu^2 p(i)}$, and $\mu = \sum_i p(i)$.

$$E = -\sum_i p(i) log\{p(i)\} \tag{9}$$

where p is the number of normalized pixel belonging to 256 bins, which were corresponding to the [0,255] grey scales of an image.

Results

To demonstrate how the proposed method worked, the intermediate results for the example image (Fig. 1) were displayed in the following subsections. This example image was randomly chosen from the 59 validation images.

Normalized chromaticity coordinates

Figure 5 shows the results after transforming the example image in Fig. 1 to normalized chromaticity coordinates. Figure 5a shows the RGB image in normalized chromaticity coordinates. R, B, and H components were proved to be most efficient from the preliminary analysis by FFSA. Figure 5b shows an RBH composite image in normalized chromaticity coordinates. In the RBH composite image, B was used instead of G, and H was used instead of B component as in an RGB image. Comparing with Fig. 1, Fig. 5a shows enhanced contrast between the fruit and background. Comparing with Fig. 5a, Fig. 5b shows more enhanced colour difference between fruit and leaves.

Potential fruit position detection using FNCC

The intermediate results from potential fruit position detection using FNCC is shown in Fig. 6. After calculating the normalized cross correlation between the template and *image_grey* from Fig. 1, the pixels with a value higher than 0.7 is shown in Fig. 6b. Although potential fruit areas were detected, there were also many other background pixels detected, because of the similarity between fruits and leaves. The green circles in Fig. 6c, d show FNCC results, which indicate potential fruit positions. The green circles in Fig. 6d indicates the potential fruit positions after removing areas smaller than 40 pixels. To remove those false positives, colour, shape and texture features were employed in the following steps.

Colour analysis

Figure 7 shows the histograms of the prebuilt data library for the chosen colour components, R, B and H, as explained in "Filter out false positives through colour analysis" section. In Fig. 7a, fruit and leaves are partially overlapped after 70, but most leaves show lower R values, which indicates R component can help with the separation of fruit and leaves. However, the histogram of background is very close to fruit, which indicates R alone is not very useful for separation of fruit and background. Figure 7b shows that fruit and background have only a small part overlapped in B histogram, which indicates that it

(a) **(b)**

Fig. 5 Results after transforming the example image to normalized chromaticity coordinates: **a** RGB image in normalized chromaticity coordinates; **b** RBH composite image in normalized chromaticity coordinates (Color figure online)

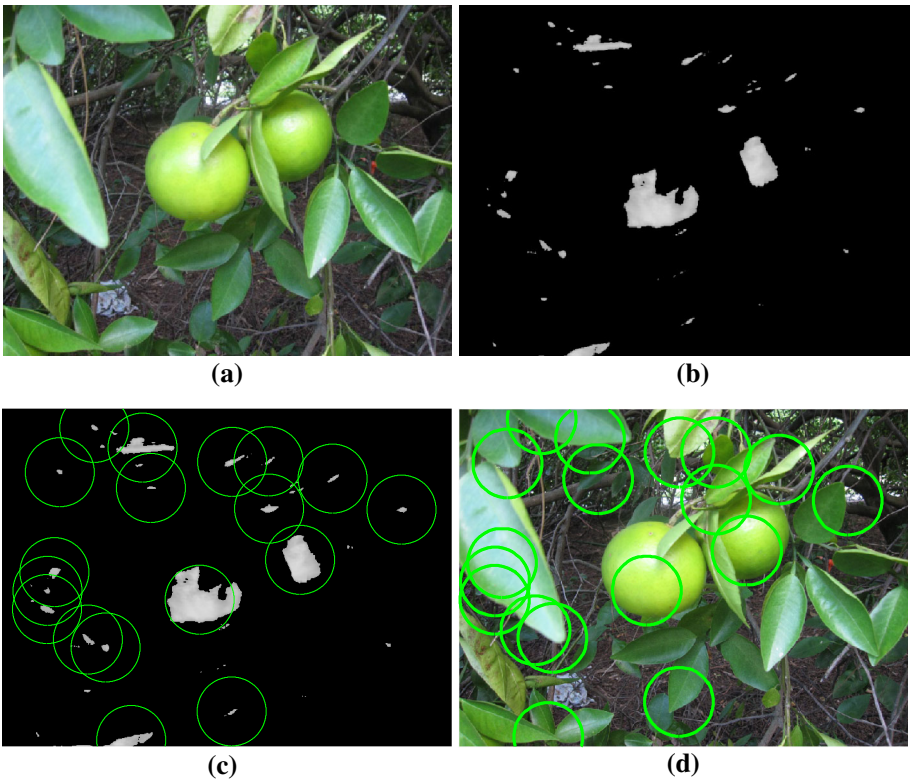

Fig. 6 An example of FNCC fruit potential position detection: **a** original image; **b** FNCC result; **c** *green circles* indicate the potential fruit positions after removing small areas; **d** FNCC result shown on the RGB image (Color figure online)

will help with the separation of fruit and background. Figure 8c indicates H value can be very helpful in both fruit-background separation and fruit-leaf separation. Overall Fig. 7 shows that colour could be a very useful feature for false positive removal after obtaining FNCC results.

A KNN classifier was used to identify the number of fruit pixels through the prebuilt fruit-background training dataset for each 30 × 30 pixel patch corresponding to the centroids in Fig. 6d. The patches with more than 450 pixels were classified as fruit. The result after false positive removal is shown in Fig. 8, in which the green circles indicate the patches being identified as fruit class. Four false positives needed to be removed further. Therefore, the shape information was utilized in the following section.

Results from CHT

Figure 9 shows intermediate results from CHT based circle detection procedure. Figure 9a shows the grey image after background removal based on the histograms shown in Fig. 7. Pixels with R value lower than 70 and with H value lower than 40 or higher than 70, were removed. Morphological opening and closing operations were applied to the background removed image, resulted in Fig. 9a. Figure 9b shows the image after 'canny' edge

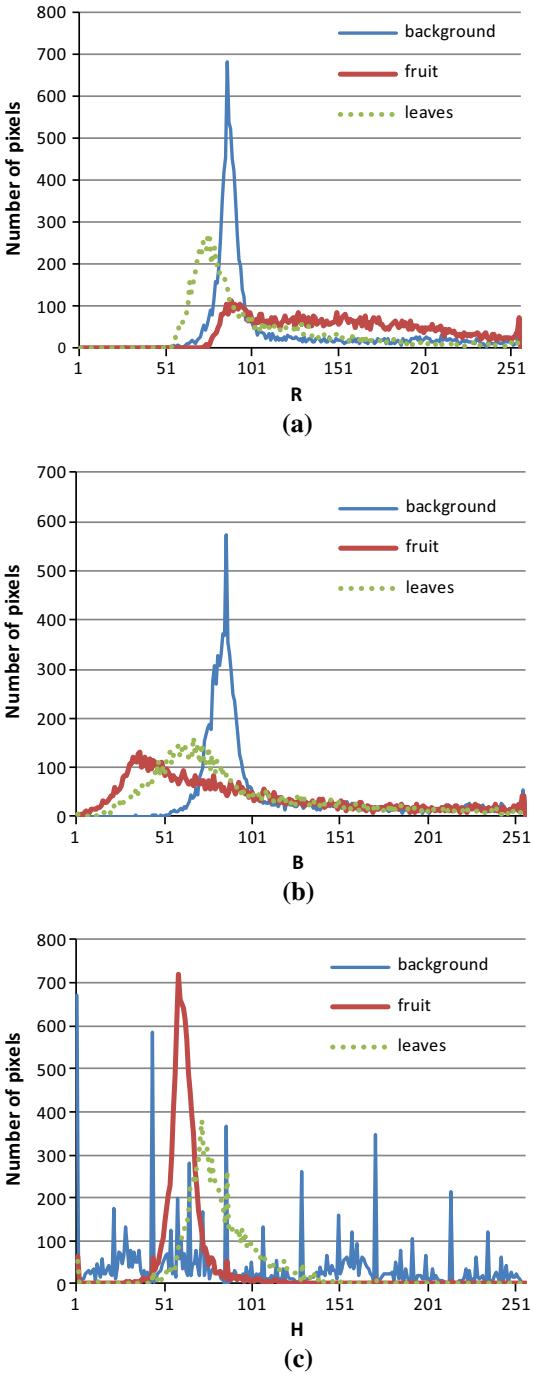

Fig. 7 Histograms of R, B, and H colour components of different objects in the training dataset: **a** histogram of R component; **b** histogram of B component; **c** histogram of H component (Color figure online)

Fig. 8 An example showing results after filtering out false positives through colour analysis (Color figure online)

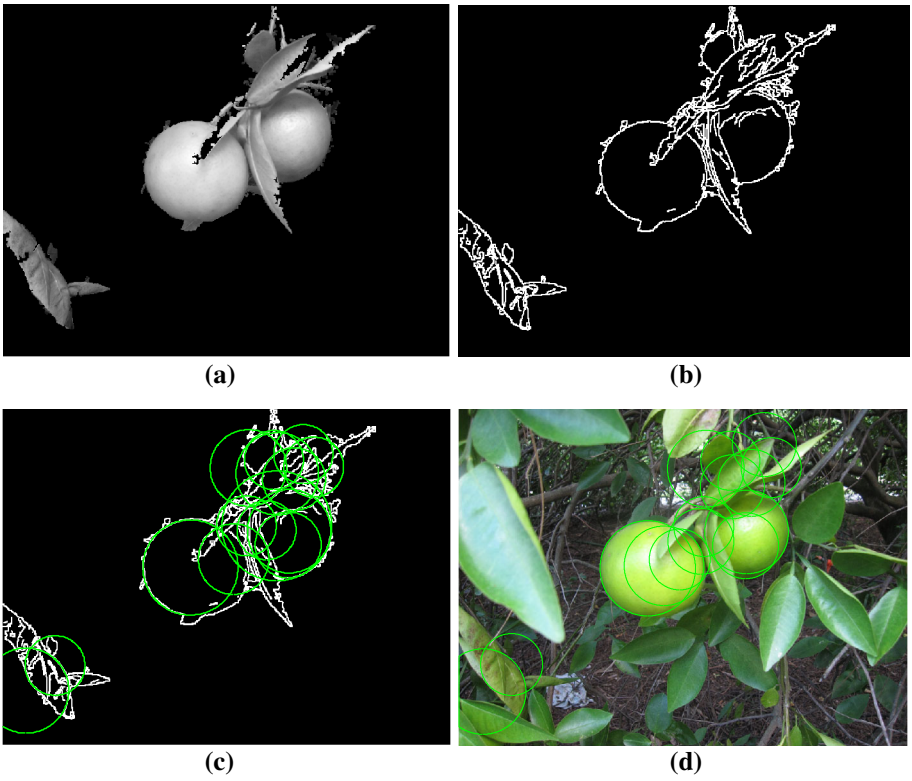

Fig. 9 Results from CHT: **a** *grey image* after background removal; **b** *edge image* of (**a**); **c** *Hough circles* on *edge image*; **d** *Hough circles* on original image (Color figure online)

detection of Fig. 9a. Figure 9c shows the circles detected by CHT, and Fig. 9d shows Hough circles overlapped on the original colour image. These results could also help with identifying the fruit size. However, since the focus of this study was to detect the number of fruits rather than estimation of the fruit size, it was not pursued further.

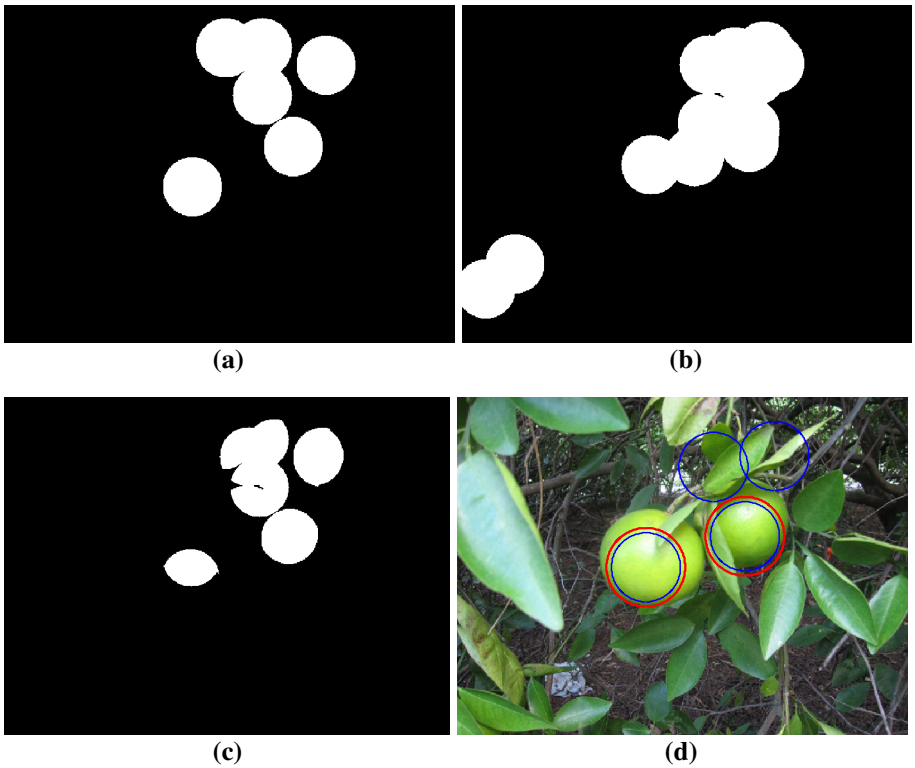

Fig. 10 Merging procedure for multiple detections and false positive removal using texture features: **a** disks centred at N2 centroids from step 2; **b** disks centred at N3 centroids from step 3; **c** intersection of (**a**) and (**b**); **d** final result: *blue colour* indicates results of merging multiple detections, and *red* indicates final detection results after false positive removal using texture (Color figure online)

Merging multiple detections and counting final number of fruit

After obtaining the detection results from steps 2 and 3, the next step was to decide how many fruit there were indeed in the image. The colour, shape and also the template features were considered for this step. Firstly a pixel disk with a radius of 60 was generated and centred at each centroid obtained from steps 2 and 3, which are shown in Fig. 10a, b. Then the intersection of Fig. 10a, b were calculated, as shown in Fig. 10c. The blue circles on the original RGB image in Fig. 10d indicate the merging result. After this step, the smoothness and entropy feature of these patches were analysed. These two features worked because the fruit patch had smoother surface, and lower entropy, in comparison to other patches with leaves and other background in them. The red circles in Fig. 10d shows the final fruit identification result with the help of texture feature after removing false positives. The fruits in the image were correctly identified.

To illustrate the performance of the proposed FNCC based green citrus detection method, Fig. 11 shows the detection results on four more example images in different imaging conditions, such as under shadow or under sunshine. The blue circles indicate results of merging multiple detections, and red circles indicate final detection results after false positive removal based on textures. These figures show successful fruit detections on

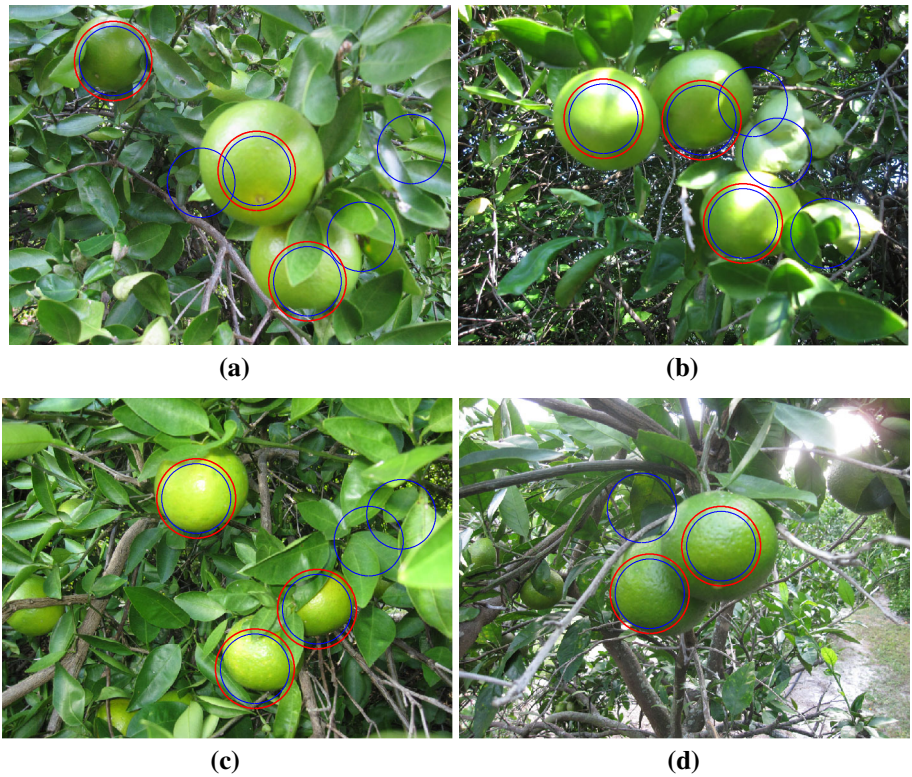

Fig. 11 More examples of fruit detection using the proposed FNCC based green citrus detection method: **a** fruits of different colour and size; **b** fruits under shadow, **c** fruits with partial occlusion; **d** fruit with brighter illumination. *Blue circles* indicate results of merging multiple detections, and *red circles* indicate final detection results after false positive removal using textures (Color figure online)

different situations: fruits of different colours and sizes (Fig. 11a), fruits in shadow and with highly saturated regions (Fig. 11b), and fruits with less than a half occluded (Fig. 11c). However, those fruits with more than a half of their edges missing or occluded with branch in the middle were hard to detect. This problem will be further studied in the future. Figure 11d shows that the fruits with brighter illumination were also correctly detected, even though the colour of the leaves and fruits were very similar and the fruits were overlapped. Through these examples, the potential of the proposed method for immature green citrus detection could be observed. It was proved that it could correctly detect fruits in different natural outdoor conditions as long as the fruits were occluded less than a half. However, for those fruits occluded more than a half, further study will be needed to detect them successfully. In contrast, Fig. 12 shows the detection results on two more example images with false fruit detection and missed fruit detection. The missed fruit detection was due to twig occlusion in Fig. 12a, and the colour similarity between the fruit and leaves in Fig. 12b. There were some other factors such as large differences in illumination conditions or visibility of fruit which could cause false or missed fruit detections using this method, which were inevitable since it is difficult for one method to resolve all the factors' influences on the results. This study focused on the general accuracy while considering as many factors as possible.

Fig. 12 More examples showing false and missed fruit detections using the proposed FNCC based method: **a** two fruits occluded by twig were missed; **b** one fruit was missed due to the colour similarity between the fruit and the leaves. *Blue circles* indicate results of merging multiple detections, and *red circles* indicate final detection results after false positive removal using textures (Color figure online)

Table 1 Fruit detection accuracy for validation dataset

	Total fruit count	Correct identification	False positive	Missed
Number of fruit	154	130	25	24
Percentage (%)	100	84.4	16.2	15.6

For all of the 59 images in the validation dataset, the same procedures were applied, and the detection results are summarized in Table 1. It should be noted here that the fruit counted were those fruits with more than a half shown in the images. There were 154 fruits in total in the validation images, and a total of 130 fruits were correctly identified, while 24 fruits were missed, and 25 false positives were introduced. The correct fruit identification accuracy was 84.4 %.

Discussions

Compared with those studies on mature fruit detection using RGB images or immature green fruit detection using multispectral or hyperspectral images, this study is much more challenging. The natural outdoor RGB images can be easily affected by the illumination change. And even with uniform illumination, there are still some big challenges from both the complicated background and the fruits themselves. On one hand, their colour was similar with that of green leaves, and on the other hand, the fruits were partially occluded by leaves, other background or even other fruits. Therefore, all possible information was needed in order to identify the green fruits correctly from the natural outdoor RGB image. In this study, three most commonly used features, i.e., colour, shape and texture were combined.

The goal of this study was to detect as many green fruits as possible in the captured image. Therefore, the first step should be detecting potential positions for the fruits, which was a very important step because those fruits lost in the first step would be irremediably lost. The results of the first step could have direct impact on the output of the subsequent

procedures. The use of FNCC made this objective possible based on the preliminary tests. This is because FNCC is a template matching based feature detection method, which utilized correlation coefficients between the template and the images. Not only the position information, but also the texture similarity was used in this method, which was exactly needed in this study.

Although colour information itself could not be used independently to identify the fruit from the image correctly, it was a good tool to help other method, especially when combing R, B and H components. After FNCC, false positives were filtered out dramatically with the help of the KNN classifier generated from the R, B, and H components of a prebuilt fruit-background dataset. What's more, before applying CHT, histograms of R, B and H components helped with the background removal. The efficiency and accuracy of CHT were improved because the edge points of the image reduced dramatically after this procedure. The texture features used were selected among several others because of the specialty of the challenges in this study. Smoothness and entropy were used in this study, because the fruit usually has a smooth surface and lower entropy compared with the complex background.

Finally, although the way of combining colour, shape and texture features in this study proved to be unique and efficient, the detection method still needs some improvement. The improvement could be obtained through improving the image acquisition time and camera parameters. For example, the illumination effect of the green citrus detection could be minimized through acquiring images close to sunset, instead of noon time, which will provide diffusive light conditions to avoid colour saturation (Linker et al. 2012). And instead of using automatic camera settings, the camera shutter parameter could be manually set to achieve more uniform illumination for the images. It is easier to detect fruit when they turn a little yellowish instead of dark green, which is helpful for distinguishing them from leaves. Also, when taking images, a constant distance could be helpful to result in more uniform fruit size.

Conclusions

FNCC based green citrus detection and counting method was proposed in this research. To use as much information as possible from the RGB image captured in natural outdoor conditions, multiple features were combined together in a unique way. The major findings in this study can be summarized as following:

(1) Normalizing RGB images to chromaticity coordinates is helpful to enhance the contrast of fruits and background.
(2) FNCC was able to identify as many potential fruit locations as possible in the first step.
(3) R, B and H colour components analysis helped in both false positive removal after FNCC and background removal before applying CHT.
(4) CHT utilized the shape information to locate the fruit, and was able to help merge multiple fruit detections and also further remove false positives.
(5) After merging multiple detections, smoothness and entropy features of the remaining patches were able to further filter out the false positives effectively.

The fruit detection accuracy of this study was 84.4 %, which indicates the potential of the proposed algorithm toward the development of an early citrus yield mapping system.

Both the algorithm and the image acquisition method will be improved and explored in the future to improve the accuracy and finally achieve the goal of early yield mapping for citrus growers.

Acknowledgments The authors would like to thank Dr. Ce Yang, Ms. Daeun Choi, Dr. Alireza Pourreza, and Dr. John Schueller at the University of Florida for their assistance in this study. The authors also would like to thank the China Scholarship Council for financial support.

References

Annamalai, P., & Lee, W. S. (2004). Identification of green citrus fruits using spectral characteristics. *ASAE Paper No. FL04-1001*. St. Joseph: ASAE.

Bansal, R., Lee, W. S., & Satish, S. (2013). Green citrus detection using fast Fourier transform (FFT) leakage. *Precision Agriculture, 14*(1), 59–70.

Brill, M. H. (2014). Definition of chromaticity coordinates. *Color Research & Application, 39*(3), 317–318.

Duda, R. O., & Hart, P. E. (1972). Use of the Hough transform to detect lies and curves in pictures. *Communications of the ACM, 15*(1), 11–15.

Duda, R. O., Hart, P. E., & Stork, D. G. (2001). Chapter 4: Nonparametric techniques. In R. O. Duda, P. E. Hart, & D. G. Stork (Eds.), *Pattern classification* (pp. 182–186). Wiley: New York.

Gong, A., Yu, J., He, Y., & Qiu, Z. (2013). Citrus yield estimation based on images processed by an android mobile phone. *Biosystems Engineering, 115*(2), 162–170.

Gonzalez, R. C., & Woods, R. E. (2002). *Digital image processing* (2nd ed.). Englewood Cliffs: Prentice Hall.

Kane, K. E., & Lee W. S. (2007). Multispectral imaging for in-field green citrus identification. *ASABE Paper No. 073025*. St. Joseph: ASABE.

Kumar, S., Ghosh, J., & Crawford, M. M. (2001). Best-bases feature extraction algorithms for classification of hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing, 39*(7), 1368–1379.

Kurtulmus, F., Lee, W. S., & Vardar, A. (2011). Green citrus detection using 'eigenfruit', color and circular Gabor texture features under natural outdoor conditions. *Computers and Electronics in Agriculture, 78*(2), 140–149.

Kurtulmus, F., Lee, W. S., & Vardar, A. (2014). Immature peach detection in colour images acquired in natural illumination conditions using statistical classifiers and neural network. *Precision Agriculture, 15*(1), 17–79.

Lewis, J. P. (1995). Fast normalized cross-correlation. *Vision Interface, 10*(1), 120–123.

Li, H., Lee, W. S., & Wang, K. (2014). Identifying blueberry fruit of different growth stages using natural outdoor color images. *Computers and Electronics in Agriculture, 106*, 91–101.

Linker, R., Cohen, O., & Naor, A. (2012). Determination of the number of green apples in RGB images recorded in orchards. *Computers and Electronics in Agriculture, 81*, 45–57.

Okamoto, H., & Lee, W. S. (2009). Green citrus detection using hyperspectral imaging. *Computers and Electronics in Agriculture, 66*, 201–208.

Pourreza, A., Pourreza, H., Abbaspour-Fard, M. H., & Sadrnia, H. (2012). Identification of nine Iranian wheat seed varieties by textural analysis with image processing. *Computers and Electronics in Agriculture, 83*, 102–108.

Sapina, R. (2001). Computing textural features based on co-occurrence matrix for infrared images. *Proceedings of 2nd international symposium on image and signal processing and analysis* (pp. 373–376). http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=938658.

Sengupta, S., & Lee, W. S. (2014). Identification and determination of the number of immature green citrus fruit in a canopy under different ambient light conditions. *Biosystems Engineering, 117*, 51–61.

Shin, J. S., Lee, W. S., & Ehsani, R. (2012). Postharvest citrus mass and size estimation using a logistic classification model and a watershed algorithm. *Biosystems Engineering, 113*, 42–53.

Stajnko, D., Lakota, M., & Hocevar, M. (2004). Estimation of number and diameter of apple fruits in an orchard during the growing season by thermal imaging. *Computers and Electronics in Agriculture, 42*, 31–42.

United States Department of Agriculture-National Agricultural Statistics Service (USDA-NASS) (2011). Forecasting Florida'' citrus production (brochure). http://www.nass.usda.gov/Statistics_by_State/Florida/Publications/Citrus/broc/1011broc.pdf.

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 Agriculture, 11*(6), 717–735.

Whitney, A. W. (1971). A direct method of nonparametric measurement selection. *IEEE Transactions on Computers, C-20*(9), 1100–1103.

Woebbecke, D. M., Meyer, G. E., Von Bargen, K., & Mortensen, D. A. (1995). Color indices for weed identification under various soil, residue, and lighting conditions. *Transactions of the American Society of Agricultural Engineers, 38*, 259–269.

Zhou, R., Damerow, L., Sun, Y., & Blanke, M. M. (2012). Using colour features of cv. 'Gala' apple fruit in an orchard in image processing to predict yield. *Precision Agriculture, 13*(5), 568–580.