An optimum method for real-time in-field detection of Huanglongbing disease using a vision sensor

Alireza Pourreza a, Won Suk Lee a,⇑, Reza Ehsani b, John K. Schueller c, Eran Raveh d

Original papers

Huanglongbing (HLB) or citrus greening is a bacterial infection which is spread by a citrus psyllid. No effective cure for this disease has been reported yet, and the HLB-infected tree will eventually die. Therefore, the infected tree must be detected and removed immediately to stop the spread of the disease. One of the symptoms of HLB is the accumulation of starch which creates blotchy mottles in an asymmetrical pattern on infected citrus leaves. These blotchy mottles symptoms may be confused with the deficiency of certain nutrients such as zinc or magnesium. We showed in a previous study that the unique capability of starch to rotate the polarization planar of light can be employed to identify the HLB-infected citrus leaves and differentiate them from zinc or magnesium deficiency. In this study, a vision sensor was developed for the purpose of real-time HLB detection for use under field conditions. The sensor included a highly sensitive monochrome camera, narrow band high power LEDs, and polarizing filters. The sensor was first tested and calibrated in a simulated field condition in a laboratory. Then, it was tested in a citrus grove. Two simple image descriptors; mean and standard deviation of gray values, were used for the purpose of classification. The results showed that the sensor clearly highlighted the starch accumulation in the HLB-infected leaf and differentiated it from visually analogous symptoms of zinc deficiency. HLB detection accuracies which ranged from 95.5% to 98.5% were achieved during the laboratory and field experiments.

1. Introduction

Huanglongbing (HLB) or citrus greening is one of the most severe infections affecting citrus production. It is caused by an Asian citrus psyllid–spread probactobacterium of the genus Candidatus Liberibacter (Albrecht and Bowman, 2008). The insect can pick up the bacteria from a citrus greening-infected tree and transfer the disease to other trees when feeding on them. In the Western Hemisphere, it was first seen in 2004 in Brazil (Texeira et al., 2005) and then in August 2005, it was reported in Florida (Halbert, 2005). Since then, the HLB disease has been reported in all citrus producing counties in Florida and has been reported at some locations in California and Texas as well. From 2004 to 2011, the Florida commercial citrus acreage and the number of trees decreased by 28% (Salois et al., 2012), and HLB was one of the major reasons for this loss. Early fruit drop which caused an average citrus production loss of 10 percent in 2012 was another consequence of HLB disease in Florida (Choi et al., 2013). Leaf yellowing that appears as blotchy mottle is one of the early HLB symptoms. Once a branch of a tree becomes infected, the disease gradually spreads through the entire tree, which will then die within 2–3 years. Severely affected trees have smaller leaves with some nutrient deficiency symptoms as well as yellow veining. Additionally, fruits from an HLB-infected tree show abnormal colors and uneven shapes and have a bitter taste. These symptoms may be used for disease diagnosis; however, they are very inaccurate, especially the older trees which suffer from multiple problems (Chung and Brlansky, 2009). Although no effective cure for this disease has been reported yet, early detection and removal of infected trees or branches were highly recommended (Buitendag and Von Broembsen, 1993) to prevent further spread of the disease.

The efficiency of an HLB diagnosis performed by several professional inspectors was evaluated in DeSoto County, Florida (Futch
The results showed that the accuracy of identifying an HLB-infected tree by visual inspection is between 47% and 59%. HLB-infected leaves contain a high level of starch accumulation (Etxeberria et al., 2009). It was shown that a leaf starch content evaluation could reveal its HLB status (Gonzalez et al., 2012). A quantitative real-time polymerase chain reaction (qrt-PCR) test is another HLB diagnosis method which can identify the HLB status with the highest accuracy (Hansen et al., 2008). Both the starch measurement and qrt-PCR tests are laboratory-based diagnostic methods and require crop scouting and leaf sampling which are expensive and time consuming. On the other hand, the HLB status of the entire grove should be observed continuously to make timely decisions and prevent huge losses. Therefore, an easy-to-use, fast, accurate, and inexpensive HLB diagnostic approach is greatly needed, especially for small growers to monitor their groves and control the spread of the disease.

A laser-induced fluorescence (LIF) spectroscopy method was employed to discriminate water-stressed and citrus canker-infected citrus leaves from healthy ones (Marcassa et al., 2006). However, their method was not able to identify citrus variegated chlorosis (CVC) disease from citrus canker disease. Later, they developed another method which was able to distinguish between mechanical interference stress and citrus canker disease stress (Belasque et al., 2008). Despite the limitations of employing LIF, some feel that this technology has much potential for citrus disease detection (Lins et al., 2009).

The potential use of green, red, and near-infrared spectral bands were determined in an HLB-infected tree identification study (Mishra et al., 2007). Later, they developed a four-band optical sensor which was able to measure the reflectance of citrus trees at 570 nm, 670 nm, 870 nm, and 970 nm. Using this sensor, they conducted multiple measurements from the same tree and obtained an error of less than 5% in the identification of HLB-infected trees (Mishra et al., 2011).

In another spectroscopy study, it was shown that the reflection of dried ground leaves in the mid-infrared band can be used to determine the HLB status of a sample with >95% accuracy (Hawkins et al., 2010). Mid-infrared spectroscopy was also employed in the classification of ground citrus leaves into three classes: HLB-infected, nutrient deficient, and healthy which resulted in an accuracy of >90% in HLB detection (Sankaran et al., 2010). In another study, near-infrared reflectance of ground citrus leaves was used to classify samples into four classes: HLB-negative, HLB-positive, nutrient deficient, and other citrus disease. The true classification rates, ranging from 92% to 99% for HLB-negative and positive samples, were acquired using a partial least squares regression model (Windham et al., 2011). Although comparatively high detection accuracies were achieved in these three studies, their methods contained sample preparation and processing, which were time-consuming and required laboratory equipment. Also these methods are impractical for in-orchard use.

Color images acquired through a digital microscope system were found to be useful in determining the HLB status of citrus leaves. Textural features were extracted from each image, and they were used to identify symptoms of HLB infection in leaf samples. An overall accuracy of 87% was achieved for classification of samples into several classes: green islands (HLB), greening blotchy mottle (HLB), normal mature leaves, young flush leaves, zinc deficiency, manganese deficiency, and iron deficiency (Kim et al., 2009). In another study, a laboratory system for laser-induced fluorescence imaging was developed in which the leaf samples were excited with a solid-state blue laser system at 473 nm. The leaf reflectance was then captured with an eight megapixel digital camera, and color descriptors were used to identify the HLB infected samples. Their results indicated a confidence level of 95% in the early stage HLB leaf identification (Pereira et al., 2011).

Airborne imagery is another approach in disease detection which has been widely used in recent years. Several HLB detection methods were conducted using airborne hyperspectral (HS) and multispectral (MS) imagery. The spectral angle mapping (SAM) classification method resulted in accuracies of 62% and 55% for MS and HS images, respectively (Li et al., 2011). It was shown later that the infected canopy had lower reflectance in the visible range and higher reflectance in the NIR range compared to a healthy canopy. Also, it was determined that the severely infected areas in the density map were easily detectable using most of the methods (Li et al., 2012b). However, the mixture tuned matched filtering (MTMF) method was proved to have a better performance compared to the SAM method in HLB detection using hyperspectral images (Kumar et al., 2012). In another study, an extended spectral angle mapping (ESAM) method was proposed for HLB detection using HS images. Accuracies of 82.6% and 86.3% were obtained for training and validation sets, respectively (Li et al., 2012a).

In our previous study (Pourreza et al., 2014), a customized image acquisition system was developed which was able to highlight an increased level of starch accumulation in HLB-symptomatic leaves. The unique starch property of rotating the polarization planar of light was used in this image acquisition system. A set of textural descriptors and step-by-step supervised classification models were used to identify HLB-positive leaves from HLB-negative and nutrient deficient samples. Although an overall accuracy of 90% was achieved using this method, some factors limited its performance in the field. The fixed automatic gain control (AGC) property of the camera made the classification process more complicated. Therefore, this study aimed to improve the performance of the prior method by implementing a completely new sensing device. Specific objectives were to (1) develop a vision sensor for real-time in-field HLB detection, (2) add on-the-go diagnosis capability to the sensor, (3) improve HLB detection accuracy, and (4) develop a simpler and more robust identification algorithm.

2. Materials and methods

2.1. Vision sensor

It was shown in the previous study (Pourreza et al., 2013) that the starch accumulation in HLB-positive leaves can rotate the polarization planar of light by 90° at 591 nm. This property was used to design the vision sensor enclosed in a wooden box (13 × 19 × 15 cm) including a camera and an illumination system (Fig. 1a). A highly sensitive monochrome camera (DMK 23G445, TheimagingSource, Bremen, Germany) with an ICX445 Sony CCD sensor was used to measure the leaf reflectance. The spectral sensitivity curve of this CCD sensor had the quantum efficiency of >90% at 591 nm which made it an appropriate option for our purpose. The camera was equipped with a wide lens (6 mm focal length) that created a diagonal field of view of 53.1° and a rotating linear polarizer and mounted inside the vision sensor housing. The very short focal length was selected to increase the depth-of-field so that more objects with different depths were in focus. Fig. 1a also shows the LED panel of the vision sensor. Ten high luminous efficiency LEDs (LED Engin, San Jose, California) at 591 nm (LZ4-00A100, 10 W) were mounted on an aluminum plate in a circular pattern. The LEDs were powered with two 12 V car batteries (24
V in total) in series and five 70 W LED drivers (RCD-48, RECOM, Brooklyn, New York) in parallel as shown in Fig. 1b. The LED panel was fixed on a side of the vision sensor, and a polarizing film (visible linear polarizing laminated film, Edmund Optics, Barrington, New Jersey) was mounted in front of it. A hole in the center of the LED panel and another one in the center of the polarizing film were cut so that there was enough room for the camera lens to come out. The direction of the camera’s linear polarizer was set to be perpendicular to the direction of the LEDs’ polarizing film as illustrated in Fig. 1a. Therefore, the camera was only able to receive the minimum reflection.

### 2.2. Data collection

A set of citrus leaf samples (‘Hamlin’ sweet orange) was collected from a grove at the Citrus Research and Education Center (CREC), University of Florida (Lake Alfred, FL) in September of 2013 by experienced HLB researchers. An experiment was conducted by acquiring images of 60 citrus leaf samples from four classes: HLB-negative (20 samples), HLB-positive (20 samples), zinc deficient HLB-negative (10 samples), and zinc deficient HLB-positive (10 samples) in a laboratory.

An in-field experiment was conducted in the CREC grove in November of 2013 in which 20 images of HLB-positive citrus trees and 10 images of HLB-negative citrus trees (the control) were acquired. Eight out of 20 samples in the HLB-positive class were also zinc deficient. The citrus trees and target leaves were located and marked by experienced researchers in the morning before the image acquisition. In order to verify the HLB status of the samples, a qrt-PCR test (Hansen et al., 2008) was performed on a total of 90 samples, including 60 of the in-lab experiment samples and one leaf sample from each image in the in-field experiment. The qrt-PCR test was conducted by acquiring images of 60 citrus leaf samples from four classes: HLB-negative (20 samples), HLB-positive (20 samples), zinc deficient HLB-negative (10 samples), and zinc deficient HLB-positive (10 samples) in a laboratory.

2.3. In-lab experiment

The lab experiment was conducted to evaluate several simulations of field imaging conditions and to determine the best settings for the sensor. Since the vision sensor had its own illumination system, the in-field experiment was conducted after sunset to prevent any interference from sunlight. Therefore, the lab experiment was conducted in a completely dark room to simulate real lighting conditions. An exposure time of 0.1 s was set for the camera because this was the shortest exposure time for capturing visually informative images without adding any gain (which increases the noise level). In order to determine the effect of the object depth on its histogram features, the images of one leaf were acquired from different distances, ranging from 50 cm to 150 cm. Then the histograms of the images taken at different depths were plotted and compared with each other. Two main histogram features including mean and standard deviation (SD) were considered for this evaluation, and the relationships between these features and the object depth was modeled.

All image acquisitions for the in-lab experiment were designed to be conducted with a fixed depth, assuming that the mean and SD features of any leaf at different depths can be computed accurately with its known depth.

In order to determine the optimum distance, three distances (60 cm, 80 cm, and 100 cm) were examined. Also, four leaf positioning conditions, including separated, adjacent, and overlapped leaves as well as the leaves on an artificial citrus tree were defined to evaluate how the leaf position in the image can influence the detection accuracy.

A circular area on each leaf was randomly selected from the symptomatic areas of HLB-positive and zinc-deficient samples, as well as a random area of HLB-negative samples. In order to select the same spot on the images of the same leaf taken from three different distances, the sizes of 177, 112, and 52 pixels were chosen for the circular areas on the images taken from 60 cm, 80 cm, and 100 cm, respectively. Then, the histograms of the two symptomatic areas and the HLB-negative regions were compared to each other to illustrate the dissimilarity of the histograms of the three different types of leaves. In order to determine whether the positioning condition of the leaves affects the identification accuracy of the symptomatic areas, a probability-based color transfer function was developed according to the histograms of symptomatic areas. In this function, three probabilities (corresponding to the three classes) were defined for each pixel value based on the histogram analysis (Eqs. (1) and (2)):

\[
P_c(i) = \frac{H_c(i)}{\sum_{c \in C} H_c(i)} \forall C = \{HLB-, HLN+, ZnDef\}
\]

\[
P_{HLB-}(i) + P_{HLB+}(i) + P_{ZnDef}(i) = 1
\]

where \(i\) is a gray value between zero and 255, \(P_c(i)\) indicates the probability that pixel value \(i\) belongs to class \(c\), and \(H_c(i)\) is the histogram value of class \(c\) for pixel value \(i\). Then, a color transfer function was developed based on these probabilities to convert the grayscale image to a red (R), green (G), and blue (B) image in which the amount of R, G, and B represents the probabilities of HLB infection, healthiness, and zinc deficiency, respectively.

![Fig. 1. Vision sensor and LED circuit: (a) a schematic of the vision sensor with dimensions in which the polarizing filters are highlighted to emphasize their relative perpendicular directions; (b) the LED circuit including 10 high power LEDs and five drivers.](image-url)
2.4. In-field experiment

In order to test the sensor in real in-field conditions, the images of citrus trees were acquired after sunset. Images were taken at an average distance of 80 cm from the trees and from the exact distance of 80 cm from the target leaves (Fig. 2). The target leaf from each image was marked and collected for a qrt-PCR test to validate its HLB status. The normalized histogram of the target leaf area in each image was obtained for further analysis.

2.5. Data analysis and classification

Two simple statistical histogram features, the mean and SD of the gray value (Eqs. (3) and (4)), were extracted from the normalized histograms \( h(i) \) (Pourreza et al., 2012) of individual leaves and leaves on the artificial tree from the lab dataset and the target leaves from field dataset.

Mean of the gray values: \( \mu = \frac{1}{s} \sum_i ih(i) \)  

SD of the gray values: \( \sigma = \sqrt{\frac{1}{s} \sum_i (i - \mu)^2 h(i)} \)

In order to evaluate the separability between the classes, a two-dimensional plot of the samples based on their means and SDs was used. Then, a maximum margin method (Bishop, 2006) was conducted to find the best divider threshold line between the classes. In this method, an objective function tries to find the optimum divider threshold which maximizes the margin between each pair of classes. A step-by-step classification model was designed based on the scatterplots of samples (Fig. 3). A support vector machine (SVM), which is also a maximum margin classifier, was trained on the scatterplots of samples (Fig. 3). A support vector machine (SVM), which is also a maximum margin classifier, was trained on the scatterplots of samples (Fig. 3). A support vector machine (SVM), which is also a maximum margin classifier, was trained on the scatterplots of samples (Fig. 3). A support vector machine (SVM), which is also a maximum margin classifier, was trained on the scatterplots of samples (Fig. 3).

3. Results

3.1. Dataset validation

The cycle threshold (CT value) in a qrt-PCR test indicates the number of required cycles for the fluorescent intensity to reach the threshold. The threshold of 33 was selected for CT values to determine the HLB status of the samples (Li et al., 2006). In other words, the samples with CT values below 33 were considered as HLB-positive leaves. Table 1 illustrates the interpretation of the CT values and HLB status for the samples used for the in-lab experiment. The CT values in all HLB-negative samples were above 33 which confirmed they were not infected. Also, all HLB-positive samples had CT values below 33 which verified their HLB infection. The CT values for half of the zinc-deficient samples were below 33 and those of the other half were above 33. Thus, there were 10 HLB-positive and 10 HLB-negative samples within the zinc deficient class.

Table 2 shows the CT values for the samples in the field experiment dataset. The CT values for ten samples were above 33, so they were categorized as the HLB-negative samples. Sample numbers one to 20 had CT values below 33, and they were considered as HLB-negative. Eight out of 20 HLB-positive samples were also zinc-deficient, and they were categorized in another subclass of HLB-positive zinc-deficient samples. Since all the zinc-deficient samples had CT values below 33, there was no zinc-deficient HLB-negative class in this dataset.

3.2. In-lab experimental results

3.2.1. The object depth effect

Fig. 4 shows the leaf sample which was used to evaluate the effect of depth and gray images acquired at different depths. In order to find the relationship between the object depth and histogram features (mean and SD), a power regression method (Gennadios et al., 1996) was employed in Excel. Figs. 5 and 6 show a line fit, regression equation, and a coefficient of determination \( R^2 \) value for mean \( \mu \) and SD \( \sigma \) based on the object depth \( d \). These curves and the very close coefficients of determination to the value of one confirmed the close relationship between the object depth and its histogram features. Therefore, these equations can be used for feature calibration as the pre-processing step for an on-the-go HLB diagnosis system when the depth information is available.

Fig. 7 shows the normalized histograms of HLB-negative, HLB-positive, and zinc-deficient symptomatic areas at three different distances: 60 cm, 80 cm, and 100 cm. The histograms of three classes were distinctive at all distances with a few overlaps; however, the range of the gray values at the distance of 100 cm was shorter than the other two distances which increased the amount of overlap between neighboring curves. At the distance of 60 cm, 92% of pixels in the zinc-deficient class were saturated (gray value \( > 255 \) which overlapped with 3% of pixels in the HLB-positive. At the distance of 80 cm, 45% of zinc deficient pixels were also saturated; however, they did not overlap with the HLB-positive pixels. The maximum range of the gray values \( \{ \max(\{x\} = 0) - \min(\{x\} = 0) \} \) was obtained at the distance of 80 cm as well. Therefore, the distance of 80 cm was chosen as the optimum distance for HLB identification in both field and lab experiments.

3.2.2. The positioning effect of the leaf

Fig. 8 shows three samples, one from each class: zinc-deficient, HLB-negative, and HLB-positive, in four leaf-positioning conditions: individual, adjacent, overlapped, and leaves on the artificial tree. The corresponding RGB images at each leaf positioning
condition were created using the color transfer function in which the green, red, and blue colors indicate the HLB-negative, HLB-positive, and zinc-deficient areas, respectively. The color transfer function was able to detect the symptomatic areas in all leaf positioning conditions.

### 3.2.3. Histogram features and classification results

The features of the mean and SD of the gray values which were extracted from the normalized histograms of the images of the citrus leaves in the dataset for the in-lab experiment are shown in Table 3. Both features for healthy samples in both leaf positioning conditions (individual leaves and leaves on the artificial tree) were generally smaller than HLB-positive and zinc-deficient samples. Also, these features were normally greater in zinc-deficient samples compared to HLB-positive samples. A comparison between the features of the same leaves in the two different leaf positioning conditions indicated that the sample images acquired on the artificial tree had smaller gray value means for 83% of samples (50 out of 60 samples) and also smaller gray value SDs for 78% of samples (47 out of 60 samples).

Fig. 9 illustrates a scatter plot of samples in four classes based on the means and SDs of the gray values for the images of the individual leaves. This plot shows a clear distinction between the HLB-negative class compared to the HLB-positive and zinc-deficient classes. Three linear thresholds were acquired from the maximum margin method and used to divide all samples into four classes of HLB-negative, HLB-positive, zinc-deficient HLB-negative, and zinc-deficient HLB-positive. The sample number 23 had the maximum SD value equal to 8.0 in the HLB-negative class (Table 3), while the minimum SD value of the classes was 11.0, which belonged to the sample number 55 in the HLB-positive class. A second linear threshold \( r_2 = -0.81\mu + 106.43 \) separated the HLB-positive samples from the zinc-deficient leaves with one misidentified HLB positive sample (#58). The last threshold was set within the zinc-deficient class \( r_3 = -0.37\mu + 15.43 \) to identify the HLB-positive samples in this class. Using the optimum threshold, only one zinc-deficient HLB-positive leaf sample (#11) was misidentified in the zinc-deficient HLB-negative class.

Fig. 10 shows a scatter plot of the samples in four classes based on the means and SDs of the gray value for the leaf images on the artificial tree. The maximum SD value in the HLB negative class...
(sample #21) was equal to 8.1, and it was smaller than the minimum SD value in the other classes (sample #55) which was equal to 8.2. However, sample #21 was misidentified as belonging to the HLB-positive class using the first linear threshold ($r_1 = /C_0$= 0.3l + 17.74) because it was considered to be a trade-off by the maximum margin method to create a more general threshold. The second threshold ($r_2 = /C_0$= 0.1l + 35.89) was set between the HLB-positive and zinc-deficient samples which resulted in one misidentified HLB-positive leaf sample (#58). The HLB-positive and -negative samples within the zinc-deficient class were closer to each other in the scatter plot of the leaves on the artificial tree compared to the individual leaves. The best possible threshold ($r_3 = -0.4l + 11.78$) was able to separate these two subclasses with two misidentified zinc-deficient HLB positive samples (#11 and #13).

As Fig. 9 suggests, except for one HLB-positive sample and one zinc-deficient HLB-positive sample, all other samples were classified correctly in this dataset. Table 4 shows the classification results confusion matrix for the data sets of individual leaves. On average, 0.33 HLB-positive samples and 0.33 zinc-deficient HLB-positive samples were misclassified in a three-fold cross-
validation. An overall accuracy of 97% was achieved using the proposed classification model and the mean and SD features. However, the purpose in this research was to detect HLB-infection. As long as the average of 0.33 HLB-positive samples was misclassified in another HLB-positive class (zinc-deficient), their HLB statuses were identified correctly. The dotted lines in Table 4 (and also Table 5) separated the HLB-positive classes from HLB-negative classes. Therefore, the overall accuracy of 98.5% was obtained when only HLB-infection was considered in the classification results evaluation.

Based on Fig. 10, one HLB-positive sample was misidentified in the zinc-deficient HLB-positive cluster and two zinc-deficient HLB-positive leaves were also misidentified as zinc-deficient HLB-negative samples. Table 5 also illustrates the confusion matrix of the classification results for the leaves of the artificial tree dataset. Analogous to the scatter plot in Fig. 10, the average of 0.33 HLB-positive samples and 0.67 zinc-deficient HLB-positive samples were misclassified in a three-fold cross validation. The overall accuracy of 95.5% was achieved in the four-class classification. Also, an overall HLB detection accuracy of 97% was achieved for leaves on the artificial tree dataset.

3.3. Field experiment results

Fig. 11 shows three samples images (one sample image per class) acquired in the field experiment. The distance between the vision sensor and the target leaf in each image (specified with a red boundary) was exactly 80 cm.

Table 6 shows the mean and SD gray value features which were extracted from the normalized histograms of the leaf images in the field experiment. Similar to the lab experimental results, the average of the feature values in the HLB-negative class was smaller than that of the HLB-positive class. Also, these features were mostly smaller for non-zinc deficient samples within the HLB-positive class.

A scatter plot of samples based on the mean and SD of the gray values' for the images from the field experiment is shown in Fig. 12. Sample #27 had the maximum SD (9.5) in the HLB-negative class which was smaller than the minimum SD in the HLB-positive class (10.5) which belonged to sample #7. However, the first linear threshold \( \sigma_1 = -0.24\mu + 23.32 \) which was obtained by the maximum margin method was set above the sample #7 (HLB-positive). In other words, one HLB-positive sample was misclassified as HLB-
negative in order to have a more general threshold. There were two subclasses of non-zinc-deficient and zinc-deficient samples within the HLB-positive class. Sample #6 in the non-zinc-deficient subclass had the maximum mean gray value of 96.1, while the minimum mean gray values in the zinc-deficient subclass was equal to 105.4 (sample #11). The second threshold ($\sigma_2 = -2.15 \mu + 263.5$) was set between the two subclasses to separate the zinc-deficient samples from the non-zinc-deficient samples within the HLB-positive superclass. Using these two simple linear thresholds, all the samples were clustered in three classes of HLB-negative, HLB-positive and HLB-positive zinc-deficient with only one misidentified sample.

Table 7 includes the classification accuracies and one misclassification error for the field dataset. The overall three-class classification accuracy of 97% was achieved using the mean and SD features and the SVM classifier.
The purpose of this study was to optimize the HLB detection performance of a previously introduced method (Pourreza et al., 2014) by developing a new vision sensor which could increase identification accuracy and decrease algorithm complexity and analysis time.

Our determination showed that the object depth had an extreme effect on the image histogram. However, there was a close relationship between the histogram features (mean and SD) and the object depth. Since these two histogram features were used for the classification purpose in this study, they could be easily calibrated and computed using the proposed regression equations when the object depth information is available. Depth cameras, such as an RGB-D (red, green, blue, and depth) camera, can measure the depth of each individual object in the image. Khoshelham and Elberink (2012) evaluated the resolution and accuracy of the depth information of a Kinect camera (Microsoft, Redmond, Washington). They determined that the Kinect depth resolution varied from 2 mm (at a distance of 1 m) to 25 mm (at a distance of 3 m). Since the maximum imaging distance in our application never goes beyond 3 m, a Kinect sensor can be used in our system to acquire the depth information with an acceptable resolution. Therefore, the mean and SD of the gray values of any leaf in a citrus tree image can be computed and used for HLB detection. This method will be used in the future in an on-the-go HLB diagnostic system. A comparison of three different distances between the sensor and leaf sample showed that the symptomatic areas in the three classes of HLB-negative, HLB-positive, and zinc-deficient samples were clearly distinguishable in all three

### Table 4
Average number of samples in the individual leaves dataset which were classified into the four classes and their corresponding classification accuracies and misclassification errors (%). The last row and column illustrate the sum of samples in the corresponding rows or columns.

<table>
<thead>
<tr>
<th>Actual class</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>HLB+</td>
<td>6.67 (95.29%)</td>
</tr>
<tr>
<td>Zn Def. HLB+</td>
<td>0.33 (4.71%)</td>
</tr>
<tr>
<td>Zn Def. HLB−</td>
<td>0.33 (8.25%)</td>
</tr>
<tr>
<td>HLB−</td>
<td>4 (100%)</td>
</tr>
<tr>
<td>Sum</td>
<td>7</td>
</tr>
</tbody>
</table>

### Table 5
Average number of samples in the leaves on the artificial tree dataset which were classified into the four classes and their corresponding classification accuracies and misclassification errors (%). The last row and column illustrate the sum of samples in the corresponding rows or columns.

<table>
<thead>
<tr>
<th>Actual class</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>HLB+</td>
<td>6.67 (92.29%)</td>
</tr>
<tr>
<td>HLB−</td>
<td>3.33 (16.75%)</td>
</tr>
<tr>
<td>Zn Def. HLB+</td>
<td>0.33 (83.25%)</td>
</tr>
<tr>
<td>Zn Def. HLB−</td>
<td>4 (100%)</td>
</tr>
<tr>
<td>Sum</td>
<td>7</td>
</tr>
</tbody>
</table>

### Table 6
The mean and SD gray value features extracted from the normalized histogram of leaf sample images from the field dataset.

<table>
<thead>
<tr>
<th>ID</th>
<th>μ</th>
<th>σ</th>
</tr>
</thead>
<tbody>
<tr>
<td>21</td>
<td>37.6</td>
<td>5.0</td>
</tr>
<tr>
<td>22</td>
<td>29.7</td>
<td>5.5</td>
</tr>
<tr>
<td>23</td>
<td>37.0</td>
<td>7.0</td>
</tr>
<tr>
<td>24</td>
<td>33.4</td>
<td>6.9</td>
</tr>
<tr>
<td>25</td>
<td>32.8</td>
<td>5.9</td>
</tr>
<tr>
<td>26</td>
<td>29.1</td>
<td>6.4</td>
</tr>
<tr>
<td>27</td>
<td>44.2</td>
<td>9.5</td>
</tr>
<tr>
<td>28</td>
<td>29.3</td>
<td>6.3</td>
</tr>
<tr>
<td>29</td>
<td>29.8</td>
<td>6.8</td>
</tr>
<tr>
<td>30</td>
<td>26.4</td>
<td>3.9</td>
</tr>
<tr>
<td>14</td>
<td>73.5</td>
<td>18.9</td>
</tr>
<tr>
<td>19</td>
<td>78.0</td>
<td>50.3</td>
</tr>
<tr>
<td>Average</td>
<td>32.9</td>
<td>6.3</td>
</tr>
</tbody>
</table>

### 4. Discussion
The purpose of this study was to optimize the HLB detection performance of a previously introduced method (Pourreza et al., 2014) by developing a new vision sensor which could increase identification accuracy and decrease algorithm complexity and analysis time.
distances; however, the maximum separation was achieved at a distance of 80 cm. The maximum range of the gray values (237) was also obtained at the distance of 80 cm, while it was found to be 219 and 229 at the distances of 60 cm and 100 cm, respectively. Additionally, the number of pixels in each pair of neighboring classes (HLB-negative and HLB-positive or HLB-positive and zinc-deficient), which had the same gray value decreased at a distance of 80 cm.

The results of identifying symptomatic areas using the color transform function confirmed that the positioning condition of the leaves did not have a significant effect on the identification accuracy. The results were slightly different for the leaves on the artificial tree condition (Fig. 8) because the distance between the leaves and the sensor was not exactly equal to 80 cm, and also the surfaces of the leaves were not precisely perpendicular to the line-of-sight of the camera. As a result, different locations on each leaf had different depths; still the symptomatic areas were distinctive for the leaves on the artificial tree condition as well. Because of the same reason, the mean and SD gray values of the leaves mostly decreased when their images were acquired on the artificial tree compared to the individual leaf positioning condition in which the distance between the leaves and the sensor was exactly equal to 80 cm (Table 3).

The mean and SD gray values of a histogram represent the overall intensity and root mean square (RMS) contrast in an image (Peli, 1990). Based on the results of this study, the mean gray values were smaller for the HLB-negative citrus leaves compared to the HLB-positive samples (Tables 3 and 6). Since the HLB-positive citrus leaves had some starch accumulation and starch has the capability to rotate the polarization planar of polarized light, the designed sensor was able to highlight the HLB symptomatic areas on the HLB-infected leaves. These highlighted HLB-positive areas had brighter pixels which caused larger mean gray values for the HLB-positive leaves. However, the portion of HLB symptomatic area varied in each infected leaf and this variation influenced the mean gray values. Therefore, although the mean gray value was a satisfying feature for identifying HLB infection, the SD of the gray value was also used for this purpose to increase the accuracy. Still, the mean gray value alone was able to differentiate the zinc-deficient leaves from non-zinc-deficient samples within the HLB-positive class for the field experiment (Fig. 12). The SD of the gray value or RMS contrast of an image indicates the dispersion of gray values from the mean. The HLB-negative (non-zinc-deficient) samples did not have any high intensity areas, so their pixel values were mostly closer to the mean and consequently, they had smaller SD. On the contrary, the HLB-negative and zinc-deficient samples had both symptomatic (=Zn or/and HLB) and non-symptomatic areas, and as a result, they had wider histogram curves and larger SD values. Accordingly, as the scatter plots in Figs. 9, 10 and 12 suggest, simple thresholds only in the SD could effectively separate HLB-negative samples from HLB-positive and zinc-deficient samples with zero error in all datasets.

Zinc deficiency develops extensive chlorosis between the veins which causes whitish yellow color in a symmetric pattern on the zinc-deficient citrus leaf. This symptom was originally brighter than the starch accumulation symptom in the HLB-positive leaves (Figs. 7 and 8) which usually caused larger mean gray values for zinc-deficient samples. Additionally, the SD values for non-zinc-deficient HLB-positive samples were usually smaller than zinc-deficient samples. A single threshold in SD for individual leaves or leaves on the artificial tree datasets (Figs. 9 and 10) could separate the HLB-positive leaves from the zinc-deficient samples with one misidentification; however, both of features were used with the maximum margin method to achieve a more general threshold.

The HLB-positive samples within the zinc-deficient class (lab experiment) had the averages of mean gray values equal to 135.1 and 118.2 for individual leaves and leaves on the artificial tree, respectively, while these averages were equal to 107.7 and 95.1 for the zinc-deficient HLB-negative samples (Table 3). However, the best possible threshold in the mean gray value (e.g. $\mu = 82.95$) would result in average clustering error rate of 35% for individual leaves and it would be worse for leaves on the artificial tree. The averages of the SD of the gray values for HLB-negative samples within the zinc-deficient class were also equal to 68.2 and 57.1 for individual leaves and leaves on the artificial tree, correspondingly, while these averages were equal to 53.4 and 48.1 for the HLB-positive samples. However the maximum clustering accuracy rate using the optimum threshold in the SD would not be over 70% for individual leaves. The coefficient of variation ($\sigma /\mu$) is a measure of relative variability which shows the dispersion of the gray values in relation to the mean (Kannan, 1981). The averages of the coefficients of variation values for the zinc-deficient HLB-positive samples were equal to 0.41 and 0.43 in individual leaves and leaves on the artificial tree datasets correspondingly, while they were equal to 0.64 and 0.60 for the zinc deficient HLB negative samples. The zinc-deficient HLB-positive samples included both HLB and zinc symptomatic areas, so they generally had smaller coefficients of variation which illustrated less gray value dispersion. The pixel values of the HLB symptomatic areas in an HLB-positive zinc-deficient sample filled the gap between the pixel values of the zinc-deficient symptomatic areas and healthy areas in the

---

Table 7

Average number of samples in the field dataset which were classified into each of the three classes and their corresponding classification accuracies and misclassification errors (%). The last row and column illustrate the sum of samples in the corresponding rows or columns.

<table>
<thead>
<tr>
<th>Actual class</th>
<th>HLB+</th>
<th>Zn Def. HLB+</th>
<th>HLB-</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prediction</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HLB+</td>
<td>3.67 (92%)</td>
<td>3 (100%)</td>
<td>0.33 (8%)</td>
<td>4</td>
</tr>
<tr>
<td>Zn Def. HLB+</td>
<td>3 (100%)</td>
<td>4 (100%)</td>
<td>4.33</td>
<td></td>
</tr>
<tr>
<td>HLB-</td>
<td>0.33 (8%)</td>
<td>4 (100%)</td>
<td>4.33</td>
<td></td>
</tr>
<tr>
<td>Sum</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>11</td>
</tr>
</tbody>
</table>

Fig. 12. Scatter plot of samples in the three classes based on their gray value mean and SD features of the normalized histograms of the target leaves from the field dataset.
histogram of the leaf, and this was the reason for smaller coefficients of variation in this subclass. Therefore, it is necessary to use both the mean and SD features in HLB-positive samples identification within the zinc-deficient class. The slopes of the threshold lines between the HLB-positive and HLB-negative samples within the zinc-deficient class were analogous (Figs. 9 and 10) in both leaves positionings conditions datasets. Sample #11 was misidentified in both lab datasets, and sample #13 was also misidentified in the leaves on the artificial tree dataset. The coefficients of variation for both of these two samples were more similar to the HLB-negative zinc-deficient samples, so their misidentifications were unavoidable.

The zinc-deficient HLB-positive samples in all datasets had higher mean gray values in all three datasets. Since there was no zinc-deficient HLB-negative sample in the field dataset, a simple threshold within a comparatively wide margin (ΔT = 9.3) could also cluster these two classes with zero error.

Between all HLB detection methods, airborne image analysis can perform the fastest diagnosis in a large area. Li et al. (2012a) obtained the best accuracy (86.3%) using airborne hyperspectral imagery; however, their method was less accurate, more expensive, and more complicated comparing to the method presented in this study. Sankaran and Ehsani (2012) also reported the best overall accuracy of higher than 94% for field HLB detection; while the vision sensor in this study was able to identify HLB infection with less than 3% error. Additionally, compared with our previous study (Pourreza et al., 2014), the HLB identification accuracy within the zinc-deficient samples increased significantly. Using only two simple statistical image descriptors in a step-by-step classification model required a computationally inexpensive analysis algorithm which is an advantage in design and development of the commercial diagnosis product.

5. Conclusions

In this study, an HLB detection method was introduced which showed improved performance in different aspects compare to the previous studies. No sample preparation such as leaf collection or grinding was required in this method and the vision sensor could detect the infection without being in contact with leaves. Only two simple features were extracted from the leaf images and used for classification purposes. This simplification decreased the analysis expense and time, and facilitated the detection process. HLB-negative samples were classified with zero error in all three datasets. Not only was the zinc deficiency accurately detected, but also the HLB infection within the zinc-deficient leaves was identified with increased accuracies. The two major components of the vision sensor were 10 high power LEDs and an inexpensive camera, and the whole sensor was assembled with less than one-thousand dollars which made it an affordable diagnosis device even for small citrus growers. Two close relationships were found between the leaf depth in the image and the gray values’ mean and SD features which were employed in leaf HLB status determination in this study. Using these two equations and a depth measurement sensor will enable this system to be used for on-the-go HLB diagnosis in future studies. Compared to our previous study, the HLB detection accuracy increased significantly in both zinc-deficient and non-zinc-deficient classes.

Acknowledgements

The authors would like to thank the Citrus Research and Development Foundation (CRDF) for supporting this research. We would also like to express our appreciation to Dr. Hamidreza Pourreza ( Ferdowsi University of Mashhad, Mashhad, Iran), and Mr. Michael Irey (United States Sugar Corporation, Clewiston, FL) for their assistance in this study.

References


Kannan, K., 1981. Percentage Coefficient of Variation. CMFRI Special Publication, 149.


