Rapid In-field Diagnosis of Huanglongbing Disease Using Computer Vision

ALIREZA POURREZA1, WON SUK LEE1*, AND ED ETXEBERRIA2

1Department of Agricultural and Biological Engineering, University of Florida, Gainesville, FL 32611

2Citrus Research and Education Center (CREC), University of Florida, Lake Alfred, FL 33850

ADDITIONAL INDEX WORDS. Citrus Greening, computer vision, disease diagnosis, HLB, starch

Citrus greening or Huanglongbing (HLB) is an extremely destructive disease, which has had an undesirable impact on the quantity and quality of the citrus production in Florida during the past few years. No effective treatment has been reported for this disease yet; therefore, rapid diagnosis and removal of the affected trees can protect the entire grove from further infection. One of the early symptoms of HLB is the accumulation of starch in affected citrus leaves which appears as an uneven yellow blotch mottled pattern on the leaf surface. This symptom looks highly analogous to some nutrient deficiency symptoms; however, starch accumulation in the HLB symptomatic area has a unique capability of rotating the polarization planar of light at a certain waveband. A computer vision approach was employed in this study to highlight the starch accumulation and differentiate it from nutrient deficiency symptoms. This system was examined under a real in-field condition to detect HLB positive, HLB negative, and zinc deficient samples. Two simple image descriptors (mean and standard deviation) were extracted from the sample images and utilized in a step-by-step classification model. The overall accuracy of 97% was achieved for identification of citrus leaves in three classes using this method.

Citrus is one of the major agricultural products in Florida with over $9 billion annual impact on the state’s economy. However, according to the citrus reference book (Salois et al., 2012), the number of citrus trees and the commercial acreage decreased by 28% in the past few years. One of the major reasons for this loss was the citrus diseases including citrus canker and citrus greening. Citrus greening, also called Huanglongbing (HLB), is a serious infection which was first confirmed in Florida in 2005 (Texeira et al., 2005). Unfortunately, no effective treatment has been reported for this disease yet. One strategy to cope with the disease involves the identification and removal of affected trees to prevent spreading of the infection throughout the entire grove. Starch accumulation is an early symptom of HLB, which causes blotchy mottle with yellowish color on the affected leaves (Etxeberria et al., 2009). The high levels of starch accumulation can be capitalized for diagnostic purposes. However, the resemblance of this yellowish symptom to some nutrient deficiencies (such as zinc and magnesium) makes it hard to detect the disease only by human observation. Futch et al. (2009) showed that the maximum accuracy of a human based inspection for HLB diagnosis was 59%.

Quantitative real-time polymerase chain reaction (qrt-PCR) is the most reliable HLB identification method which should be conducted in the laboratory (Hansen et al., 2008). Measuring the starch content of a citrus leaf has also been used as an HLB diagnosis method (Gonzalez et al., 2012). Although these are comparatively accurate approaches, they are time consuming, laborious and cannot be employed for continuous and rapid grove monitoring.

Spectroscopy has been widely employed by researchers for early HLB identification (Mishra et al., 2012, Sankaran et al., 2010, Windham et al., 2011). For instance, Pereira et al. (2011) achieved an accuracy of 95% for identification of HLB affected leaves from healthy samples using a laser-induced florescence imaging system. Airborne imagery was also adopted for finding the HLB affected areas in citrus groves (Garcia-Ruiz et al., 2013, Kumar et al., 2012, Li et al., 2012). It was shown that up to 86% of HLB detection accuracy could be achieved by analyzing the aerial hyperspectral photographs (Li et al., 2014).

It was previously shown that a customized imaging sensor could highlight the HLB symptomatic areas on an affected leaf with good accuracy (Pourreza et al., 2014). The objective of this study was to develop a vision sensor to improve the HLB detection accuracy of the previously introduced approach especially for the zinc deficient samples.

Materials and Methods

The unique starch capability of rotating the polarization planar of light was shown in the previous study (Pourreza et al., 2014). This feature was used in designing a new vision sensor which was able to highlight the HLB symptoms on an affected citrus leaf. The vision sensor included 10 high power narrow band LEDs at 591 nm (LZ4-00A100, LED Engin, San Jose, CA) and a sensitive camera (DMK 23G445, TheImagingSource, Bremen, Germany) equipped with a linear polarizer (Fig. 1). A polarizing film was mounted in front of the LEDs with a perpendicular direction to the cameras’ polarizer so the camera receives only the minimum reflection at 591 nm.

In-field data collection was conducted at the Citrus Research and Education Center (CREC), University of Florida, Lake Alfred, Florida. Since the vision sensor had a customized illumination

*Corresponding author; phone: (352) 392-1864; email: wslee@ufl.edu
system, the image acquisition was performed after sunset to prevent the interference of other light sources. Images were acquired from 30 citrus trees (‘Hamlin’ sweet orange) at a distance of 80 cm to a target leaf in each image. The dataset included 12 HLB-positive, 8 HLB-positive zinc-deficient and 10 HLB-negative samples. The HLB statuses of the target leaves were confirmed by a qrt-PCR test at the United States Sugar Corporation (USSC), Technical Operations, Southern Gardens (Clewiston, FL).

The gray values’ mean and standard deviation (SD) features were extracted from the target leaf areas in the images and were used to illustrate the samples in a two dimensional plot. In order to find the best divider threshold between the three classes, maximum margin method (Bishop, 2006) was employed. A classification model was designed based on the samples distribution in the two dimensional plot in which the classification was conducted in two steps (Fig. 2). At the first step, HLB-negative samples were separated from the rest of dataset, and at the second step zinc-deficient samples were identified within the HLB-positive super class. A support vector machine (SVM) classifier was employed for both steps of the classification model. A k-fold cross validation method was used in which the entire dataset was randomly divided into three folds and the classifier was trained with two folds, while the third fold was used as the validation set. These steps were repeated fifty times and the averages of classification and misclassification rates were computed as the accuracy results.

MATLAB (version R2011a, MathWorks, Natick, MA) software was used for feature extraction and classification. Also Excel (Microsoft Office, Microsoft, Redmond, Washington) software was used to visualize the samples in a two dimensional plot.

Results and Discussion

Figure 3 shows three examples of citrus tree images from three classes and the target leaf which is specified with a red contour. Target leaf in each image was exactly at the distance of 80 cm to the sensor. The cycle thresholds (CT values in a qrt-PCR test) of the target leaf area in each image are listed in Table 1. CT value specifies the number of cycles that the fluorescent intensity needs to attain the threshold. Li et al. (2006) suggested that samples with CT values below 33 should be considered as HLB affected. The CT values in Table 1 validated our dataset. Table 1 also includes the mean and SD features extracted from the target leaf area in each image. In average HLB-negative samples had smaller means and SDs compared to HLB-positive samples. The starch contents in HLB positive samples exceeded the normal level, and the vision sensor highlighted those areas with brighter pixels. That was the reason for greater means in HLB-positive samples. The HLB-positive samples also included some asymptomatic areas with no starch accumulation and smaller pixel values which resulted in wider histogram curve and greater SD values compared to HLB-negative samples.

The third class of zinc-deficient HLB-positive samples had the greatest average mean and SD. The samples in this class included asymptomatic, HLB symptomatic, and zinc deficiency symptomatic areas. Zinc deficiency creates whitish yellow color symptoms caused by extensive chlorosis developed between the veins. The pixels belonging to the zinc deficiency symptoms had

Fig. 1. Vision sensor including a camera, LEDs, polarizing filters, aluminum plate, and a wooden housing.

Fig. 2. The two-step classification model; the input samples at each step were classified in two classes.

Fig. 3. Image samples and target leaves in three classes of HLB-negative, HLB-positive, and zinc-deficient HLB-positive.
higher intensities than HLB symptomatic areas which generated greater means for zinc-deficient samples.

Figure 4 illustrates the samples in a two dimensional scatter plot and the thresholds between the three classes achieved by the maximum margin method. According to the recommended thresholds, only one HLB-positive sample (# 7) was mis-clustered in HLB-negative cluster. Sample # 7 had the SD value of 10.5 which is greater than the maximum SD value in the HLB-negative class (equal to 9.5) belonging to sample # 27. Thus, a single threshold in SD values could separate HLB-positive and HLB-negative samples with no error. However, the maximum margin method tries to maximize the margin between two classes to achieve a more general threshold. That is to say one sample was mis-clustered as a penalty to obtain a more general threshold. The HLB-positive and zinc-deficient HLB-positive classes were well separated by the suggested threshold. Sample # 6 in HLB-positive class had the minimum mean in this class equal to 96.1 while the minimum mean in zinc-deficient HLB-positive class (belonged to sample # 11) was equal to 105.4. Again a simple threshold in mean values could separate the two classes with zero error but the recommended threshold provides a wider divider margin.

The classification accuracies and misclassification errors are shown in Fig. 5. All HLB-negative and zinc-deficient HLB-positive samples were classified correctly. Similar to the clustering results in Fig. 4, only one HLB-positive sample was misclassified in HLB-negative class which resulted in an 8% misclassification error. Overall accuracy of 97% was achieved for a 3-class classification using mean and SD features and a 2-step classification model.

Comparing with other methods, the introduced method has some advantages which make it a suitable option for real-time applications. This vision sensor was developed with inexpensive components and can be assembled with less than a thousand dollars. As well as its capability to detect the HLB infection, the sensor was able to differentiate the zinc deficiency from HLB infection within the zinc deficient samples. Additionally, this sensor can be used for an “on-the-go diagnosis purpose”. While mounted on a vehicle, it can autonomously acquire the tree images and assign them with the coordinates obtained by a DGPS receiver. Finally, a post-processing software can generate a map of HLB affected areas in the grove which will be used for a better field management and results in an increased profit.

**Conclusions**

A new prototype of the HLB diagnosis vision sensor was introduced in this study and its accuracy was evaluated. No leaf

Table 1. The CT values, gray values, mean, and SD features of target leaves in three classes.

<table>
<thead>
<tr>
<th>ID</th>
<th>CT Value</th>
<th>Mean</th>
<th>SD</th>
<th>ID</th>
<th>CT Value</th>
<th>Mean</th>
<th>SD</th>
<th>ID</th>
<th>CT Value</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>21</td>
<td>40.0</td>
<td>37.6</td>
<td>5.0</td>
<td>1</td>
<td>21.9</td>
<td>73.7</td>
<td>22.3</td>
<td>9</td>
<td>20.7</td>
<td>119.8</td>
<td>38.9</td>
</tr>
<tr>
<td>22</td>
<td>40.0</td>
<td>29.7</td>
<td>5.5</td>
<td>2</td>
<td>24.1</td>
<td>39.0</td>
<td>22.6</td>
<td>10</td>
<td>26.0</td>
<td>126.7</td>
<td>27.4</td>
</tr>
<tr>
<td>23</td>
<td>36.0</td>
<td>37.0</td>
<td>7.0</td>
<td>3</td>
<td>24.3</td>
<td>89.3</td>
<td>38.9</td>
<td>11</td>
<td>24.2</td>
<td>105.4</td>
<td>38.5</td>
</tr>
<tr>
<td>24</td>
<td>37.2</td>
<td>33.4</td>
<td>6.9</td>
<td>4</td>
<td>21.0</td>
<td>45.1</td>
<td>22.1</td>
<td>15</td>
<td>23.2</td>
<td>159.2</td>
<td>69.3</td>
</tr>
<tr>
<td>25</td>
<td>40.0</td>
<td>32.8</td>
<td>5.9</td>
<td>5</td>
<td>19.5</td>
<td>57.4</td>
<td>16.5</td>
<td>16</td>
<td>26.5</td>
<td>122.1</td>
<td>37.4</td>
</tr>
<tr>
<td>26</td>
<td>36.2</td>
<td>29.1</td>
<td>6.4</td>
<td>6</td>
<td>22.4</td>
<td>96.1</td>
<td>24.5</td>
<td>17</td>
<td>23.3</td>
<td>211.6</td>
<td>58.1</td>
</tr>
<tr>
<td>27</td>
<td>40.0</td>
<td>44.2</td>
<td>9.5</td>
<td>7</td>
<td>22.9</td>
<td>40.5</td>
<td>10.5</td>
<td>18</td>
<td>22.4</td>
<td>126.6</td>
<td>72.0</td>
</tr>
<tr>
<td>28</td>
<td>40.0</td>
<td>29.3</td>
<td>6.3</td>
<td>8</td>
<td>23.0</td>
<td>86.5</td>
<td>37.6</td>
<td>20</td>
<td>21.5</td>
<td>127.1</td>
<td>40.2</td>
</tr>
<tr>
<td>29</td>
<td>40.0</td>
<td>29.8</td>
<td>6.8</td>
<td>9</td>
<td>24.3</td>
<td>94.2</td>
<td>28.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>40.0</td>
<td>26.4</td>
<td>3.9</td>
<td>10</td>
<td>26.1</td>
<td>69.1</td>
<td>27.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>11</td>
<td>22.6</td>
<td>73.5</td>
<td>18.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>12</td>
<td>21.4</td>
<td>78.0</td>
<td>50.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Average: 38.9  32.9  6.3  Average: 22.8  70.2  26.7  Average: 23.5  137.3  47.7

Fig. 4. Two dimensional scatter plot of samples and the thresholds between the three classes. The samples which are specified with a circle were selected by the maximum margin method as the support vectors.

Fig. 5. Classification rates and misclassification errors (%) results for samples in three classes.
collection or sample preparation was required in this method and it can be adopted for an on-the-go application. Two simple image descriptors, mean and SD, were extracted from the leaf area and were employed to train the SVM classifier. It was shown that in some cases, one descriptor alone could separate the classes with zero error. This improvement was achieved due to the customized design of the narrow band illuminations system and the proper use of polarizing filters. The results of this study showed an enhanced accuracy and ease of use comparing to our previous study.

**Literature Cited**


