# Detection of citrus Huanglongbing based on image feature extraction and two-stage BPNN modeling

Deng Xiaoling[1,2], Yubin Lan[1,3*], Xing Xiaqiong[2], Mei Huilan[1,3], Liu Jiakai[2], Hong Tiansheng[3]

(1. *International Lab of Agricultural Aviation Pesticides Spraying Technology, Guangzhou 510642, China*;
2. *College of Electronical Engineering, South China Agricultural University, Guangzhou 510642, China*;
3. *College of Engineering, South China Agricultural University, Guangzhou 510642, China*)

**Abstract:** Citrus Huanglongbing (HLB), which is spread by the citrus psyllid, is the most destructive disease of citrus industry. While no effective cure for the disease has been reported, detection and removal of infected trees can prevent spreading. Symptoms indicative of HLB can be present in both HLB-positive trees and HLB-negative trees, making identification of infected trees difficult. A detection method for citrus HLB based on image feature extraction and two-stage back propagation neural network (BPNN) modeling was investigated in this research. The identification method for eight different classes including healthy, HLB and non-HLB symptoms was studied. Thirty-four statistical features including color and texture were extracted for each leaf sample, following the two-stage BPNN to model and identify HLB-positive leaves from HLB-negative leaves. The discrimination accuracy can reach approximately 92% which shows that this method based on visual image processing can perform well in detecting citrus HLB.

**Keywords:** citrus leaf, Huanglongbing, texture and color features, feature extraction, two-stage back propagation neural network
**DOI:** 10.3965/j.ijabe.20160906.1895

**Citation:** Deng X L, Lan Y B, Xing X Q, Mei H L, Liu J K, Hong T S. Citrus Huanglongbing detection based on image feature extraction and two-stage back propagation neural network modeling. Int J Agric & Biol Eng, 2016; 9(6): 20－26.

## 1 Introduction

Citrus Huanglongbing (HLB), also named as citrus yellow shoot disease, is thought as originating from China in the early 1900s. The disease in Asian areas is primarily caused by bacterium and spread by Asian citrus psyllid, thriving on young citrus leaves[1]. The bacteria itself is not harmful to humans, but can cause an obstruction in the flow of nutrients in citrus trees. This in turn causes limb dieback, diminishes production and ultimately kills the tree[2]. As the health of the tree declines, there is an increase in early fruit abscission with the fruit that is produced being bitter and misshapen while remaining green and small[3]. As there is currently no medicine for this disease, prevention is critical to the maintaining of quality citrus production. The most common method for diagnosing HLB infected trees is human visual identification. However, as this method is limited to the personal knowledge and experience of the person performing the inspections, it typically leads to low recognition accuracy. More reliable laboratory diagnostic methods of HLB have been reported[4], including electron microscopy, Thin Layer

**Received date:** 2015-04-25 **Accepted date:** 2016-06-09
**Biographies: Deng Xiaoling**, PhD, Associate Professor, research interests: agricultural information and aviation application, Email: dengxl@scau.edu.cn; **Xing Xiaqiong**, Bachelor, research interests: signal processing, Email: 1041900290@qq.com; **Mei Huilan**, PhD, research interests: digital signal processing, Email: 63629353@qq.com; **Liu Jiakai**, Bachelor, research interests: signal processing, Email: 1134018379@qq.com; **Hong Tiansheng**, PhD, Professor, research interests: agricultural information and machinery application, Email: tshong@scau.edu.cn.
***Corresponding author: Yubin Lan**, PhD, Professor, research interests: agricultural aviation application. Mailing address: International Laboratory of Agricultural Aviation Pesticide Spraying Technology (ILAAPST), South China Agricultural University, Guangzhou 510642, China. Email: ylan@scau.edu.cn.

Chromatography (TLC)[5], and Polymerase Chain Reaction (PCR)[6], et al. Machine vision identification has also been attempted. Pereira et al.[7] analyzed the color of laser-induced fluorescence images, and observed that the colors of HLB infected leaves and healthy leaves could be distinguished in eight color channels. Deng et al.[8] used a Gaussian mixture density (GMD) to extract the leaf object from the citrus visible image, followed by a feature extraction and recognition of HLB in the leaf based on scalable vocabulary tree. Color Co-occurrence Matrix (CCM) was used to distinguish different symptoms of citrus disease[9], yet by ignoring the influence of the gray scale, it could not achieve a high classification accuracy. A sensing system based on narrow-band imaging and polarizing filters was developed to discriminate between HLB-symptomatic and healthy leaves[10]. However, the method considered a few of the citrus symptoms. Methods using hyperspectral imaging technology have also been proposed. Mei et al.[11] adopted a partial least squares-discriminate analysis (PLS-DA) method to analyze the hyperspectral image data of citrus leaves for early detection of citrus HLB. Yet the accuracy of diagnosis still needs to be improved. Deng et al.[12] adopted a Minimum Noise Fraction (MNF) Rotation for dimensionality reduction of hyperspectral data, created training set of pure pixels using Pixel Purity Index (PPI), and identified HLB with Fisher discrimination. A method based on principal component analysis and BPNN based on hyperspectral images was used to diagnose and classify citrus HLB with an accuracy of over 90%[13]. Airborne multispectral and hyperspectral imaging methods have also been adopted in detection of HLB[14-17]. Kumar et al.[1] utilized airborne multispectral information and hyperspectral imaging to perform citrus HLB detection; the highest recognition accuracy was about 86%. However, to distinguish HLB-infected trees from other nutrient-deficiency trees was proved difficult. Using the reflectance of healthy and HLB canopies in both the visible and near-infrared ranges was able to detect most of the severely infected areas. The detection accuracies ranged from 43% to 95%, depending on the ground truth accuracy and the classifiers adopted.

Thus, while a number of HLB detection methods have been developed and evaluated, the results were not satisfactory as a result of long detection time and low accuracy. The overall objective of this research was to develop an effective feature extraction method to reveal the HLB heterogeneity from visual leaf image and distinguish citrus HLB infected tress from other non-HLB (HLB-negative) trees with similar symptoms. The method presented involves four major processes:

1) RGB images of orange leaves collecting, eight different symptoms (i.e., floral leaf, blocky mottled, uniform yellowing, normal mature, root rot yellowing, phytotoxic yellowing, edge yellowing and other non-HLB yellowing) capturing.

2) Digital images of leaves preprocessing;

3) Texture and color features extracting to reveal HLB heterogeneity;

4) Two-stage BPNN modeling and classification of citrus HLB.

The framework of citrus disease recognition is shown in Figure 1.

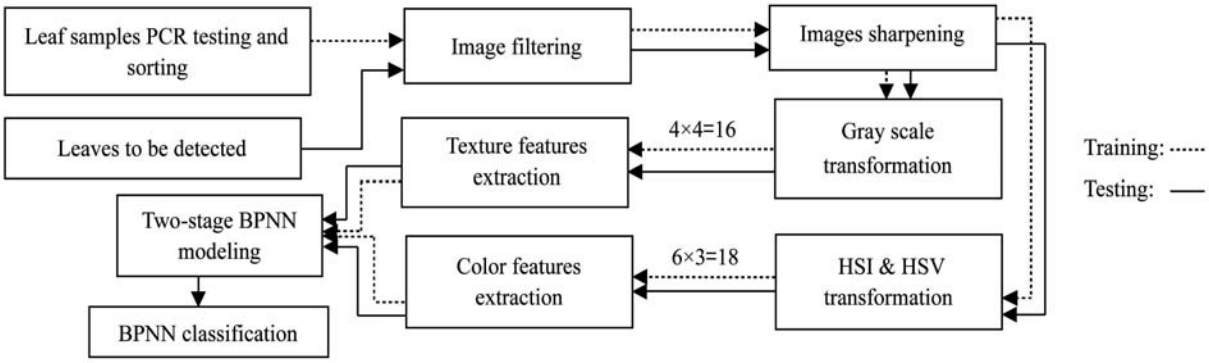
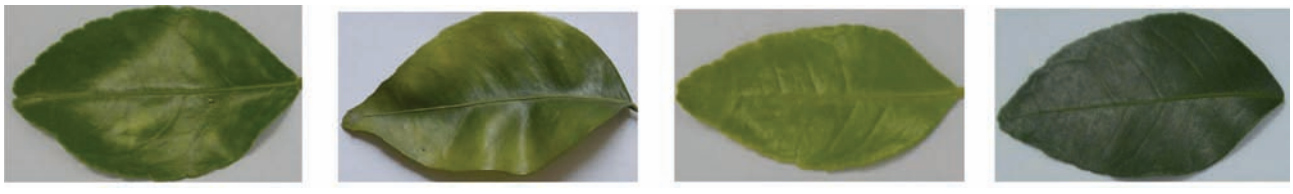



Figure 1 Framework of citrus HLB recognition

## 2 Materials and methods

### 2.1 Materials

The leaf samples used in this research were collected from two sugar sand orange orchards in Qingyuan and Huizhou located in Guangdong Province, China, in September 2013 and 2014. The citrus leaf images were acquired using a NIKON D600 camera under natural light. The dimensions of each leaf sample image were resized to no larger than 1260×720. Eight different classes of citrus leaves were selected for the experiment, with nearly 100 samples in each class being collected and tested by both PCR detection and observation by experienced experts. The classes of leaf sample include the following: floral leaf, blocky mottled, uniform yellowing, normal mature, root rot yellowing, phytotoxic yellowing and edge yellowing. Images of leaf samples are shown in Figure 2. The first three classes (a-c) represent the typical symptoms of citrus HLB, the normal mature class (d) represents healthy trees, and the final four classes (e-h) represents other similar symptoms from non-HLB infected trees. Each class had distinctive features on the leaf surface that represented the particular symptoms (Figure 2). The blocky mottled leaf was covered with irregular patches. The uniform yellowing leaf was covered with a single yellow color on the whole surface. The root rot yellowing leaf presented an abnormal color, especially on the petiole. Given that the goal of this work was to distinguish HLB symptoms from non-HLB symptoms, the last four classes (e-h) of citrus leaves were classified as a single class - the leaves of non-HLB trees. The combination of non-HLB symptoms into a single class results in five leaf symptom classes that needed to be trained and distinguished.



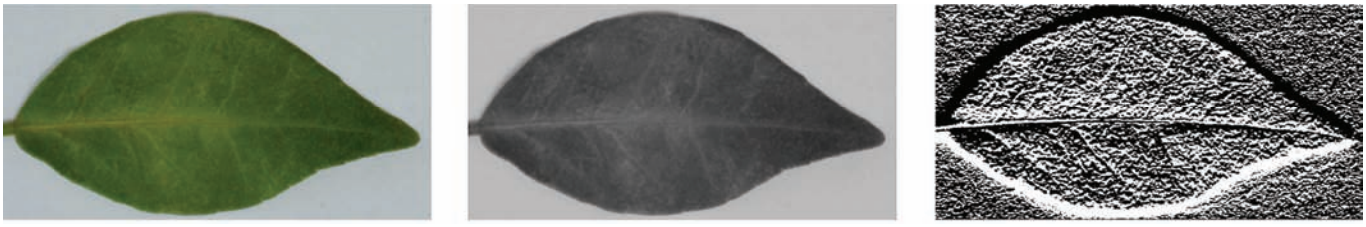
a. Floral leaf b. Blocky mottled c. Uniform yellowing d. Normal mature

e. Root rot yellowing f. Phytotoxic yellowing g. Edge yellowing h. Other non-HLB yellowing

Figure 2 Leaf samples

### 2.2 Preprocessing of the samples

As photographing conditions were different, the preprocessing of image was required. Also, any differential operator highlights the noise if not been properly filtered. Therefore the preprocessing system used in this study first conducted filter processing followed by image sharpening processing. As median filter can preserve edges while removing noise[18], a median filter was processed next. A Sobel operator was used in image sharpening considering that it is relatively effective in fringe detection[19] by allowing for clear image outlines all while restraining noise and enhancing the contrast of the gray scale. Post-preprocessed images of an example floral leaf are shown in Figure 3. Matlab R2010b software was used for image preprocessing and two-stage BPNN modeling.



a. Non-preprocessing b. Median filter c. Sobel operator

Figure 3 Preprocessing images

### 2.3 Texture features extraction

The Gray level co-occurrence matrix (GLCM) is a statistical method for determining a shape by statistically sampling gray-level gradients across certain distances and directions within an image[20]. It is a tabulation of how often different combinations of pixel brightness values (gray levels) occur in an image. The GLCM functions characterize the texture of an image by calculating how often pairs of pixel with specific values and in a specified spatial relationship occur in an image, creating a GLCM, and then extracting statistical measures from this matrix. Given a random point ($x$, $y$) and its adjacent point ($x+a$, $y+b$) of an image ($M\times N$), the point pair's gray level is notated as (g1, g2). The probability of two pixel gray scales occurring together will transform the spatial coordinate ($x$, $y$) into describing the 'gray level pair' (g1, g2) and form the GLCM.

In order to make the classifications more accurate, the GCLM was applied and constructed in four directions including 0º, 45º, 90º and 135º[21]. The non-normalized GLCM is defined by:

$$P(g_1, g_2, \theta) = \sum_{x=1}^{M}\sum_{y=1}^{N} C\{(I(x,y)=g_1)\Lambda(I(x\pm db, y\mp da)=g_2)\} \quad (1)$$

where, $\theta$ is one of the four directions above. When $a$=1, $b$=0, the pixel pair is horizontal, i.e., 0º scanning; when $a$=0, $b$=1, the pixel pair will be vertical, i.e., 90º scanning; when $a$=1, $b$=1, the pixel pair will be right diagonal, i.e., 45º scanning; when $a$=−1, $b$=1, the pixel pair will be left diagonal, i.e., 135º scanning. If the argument is true, $C\{\cdot\}$=1, otherwise, $C\{\cdot\}$=0. The ± and ∓ signs mean that each pixel pair is counted twice: once forward and once backward.

To obtain standard uniform parameters, the probability parameter of the GLCM was also normalized through dividing by the total number of counted pixel pairs. After normalizing the GCLM, four statistical texture features would be extracted: energy, entropy, inertia and dependency, intensity texture feature equations are presented in Table 1. Therefore, including every eigenvalue extracted in four directions of 0º, 45º, 90º and 135º, there would be totally 16 characteristic vectors representing each texture feature.

**Table 1 Intensity texture feature equations**

| Feature | Description | Equation | |
|---|---|---|---|
| $Q_1$ | Energy | $Q_1 = \sum_{g_1}\sum_{g_2}[p(g_1,g_2)]^2$ | (2) |
| $Q_2$ | Entropy | $Q_2 = \sum_{g_1}\sum_{g_2} p(g_1,g_2)\lg p((g_1,g_2)$ | (3) |
| $Q_3$ | Inertia | $Q_3 = \sum_{g_1}\sum_{g_2} k^2 p(g_1,g_2)^2 \cdot k = \lvert g_1 - g_2 \rvert$ | (4) |
| $Q_4$ | Dependency | $Q_4 = \dfrac{\sum_{g_1}\sum_{g_2} g_1 g_2 p(g_1,g_2) - \mu_x\mu_y}{\delta_x\delta_y}$ | (5) |
| | | $\mu_x = \sum_{g_1} g_1 \sum_{g_2} p(g_1,g_2),\ \mu_y = \sum_{g_2} g_2 \sum_{g_1} p(g_1,g_2)$, $\delta_x^2 = \sum_{g_1}(g_1-\mu_x)^2\sum_{g_2} p(g_1,g_2),\ \delta_y^2 = \sum_{g_2}(g_2-\mu_y)^2\sum_{g_1} p(g_1,g_2)$ | |

### 2.4 Color feature extraction

To improve the efficiency and precision of data processing, an appropriate color space should be chosen according to the application environment[22]. Common color spaces in digital image processing include RGB, HSI, HSV etc., which have different characteristics and scopes of application. Though RGB is the most common color space, it is typically used as a visual display of the image. HSI color space attempts to produce a more intuitive representation of color, greatly reducing the workload of image analysis and processing by allowing for a large number of algorithms to be conveniently performed. HSV color space enables the saturation of a greater dynamic range, which is best for hue and saturation processing. For these reasons, HSV and HSI color spaces were selected in this study. Thus, HSI and HSV transformations were performed and three statistical features including the first rank moment (mean value), the second rank moment (standard deviation), and the third rank moment were extracted. The intensity color feature equations for these three features are presented in Table 2. Three rank moments were extracted in each channel of six color channels resulting in 18 eigenvalues. Therefore, including the texture and color features discussed in the previous section, totally 34 characteristics for each sample were included in this study.

**Table 2 Intensity color feature equations**

| Feature | Description | Equation | |
|---|---|---|---|
| $Q_5$ | The first rank moment | $\bar{X} = \dfrac{\sum_{i=1}^{N} X_i}{N}\ (i=1,2,3,...,N)$ | (6) |
| $Q_6$ | The second rank moment | $\sigma = \sqrt{\dfrac{\sum_{i=1}^{N}(X_i-\bar{X})^2}{N}}\ (i=1,2,3,...,N)$ | (7) |
| $Q_7$ | The third rank moment | $\kappa = \sqrt[3]{\dfrac{\sum_{i=1}^{N}(X_i-\bar{X})^3}{N}}\ (i=1,2,3,...,N)$ | (8) |

### 2.5 Two-stage BPNN for modeling and classification

The BPNN with 34 input nodes and five output nodes was adopted for modeling and classification to allow for 34 feature vectors and five different classes. In this study, a two-stage BPNN was constructed. The recognition process is shown in Figure 4. As there were different kinds of non-HLB leaves, a prediction model could not be directly constructed. In this research a first stage neural network (marked as A in Figure 4) was built to train the datasets of the four HLB leaf classes (uniform yellowing, floral, normal mature and mottled) showed in Figure 2. In the second stage, four sub-neural networks were built to train the datasets for uniform yellowing & non-HLB, floral & non-HLB, normal mature & non-HLB and mottled & non-HLB. When the sub-neural networks returned the best results, they were saved and marked as B, C, D and E. A minimum MSE (mean squared error) criterion was used to determine the best sub-neural network training. Output 1 (Figure 4) represents the uniform yellowing symptom, output 2 represents the floral symptom, output 3 stands for normal mature, output 4 stands for blocky mottled symptom and output 5 stands for yellowing leaves of non-HLB symptom.

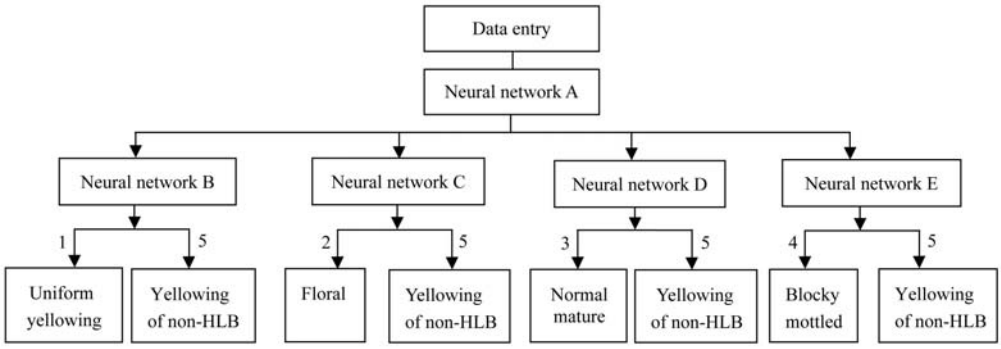



Figure 4 Recognition process of the two-stage BPNN

## 3 Results and discussion

As mentioned above, based on results from PCR testing and experimental judgments from agricultural experts, the leaves were classified into five classes of floral, blocky mottled, uniform yellowing, normal mature and yellowing of non-HLB. Thirty-four vectors, including texture and color features, were extracted requiring an approximate total time of 160 s per sample. From the total 500 groups of feature data, 250 groups were selected randomly for the training of the neural network, the remaining data were used for the testing of the resulting best fit algorithms. Based on the structure of the two-stage BPNN, the 5th class, representing yellowing of non-HLB leaves, might be incorrectly assigned to any of the other four classes in the first stage. Therefore, the second stage sub neural networks (B-E) must be different between each of the HLB positive leaf classes and the non-HLB leaf 5th class showing similar symptoms. For example, if the 5th class was incorrectly assigned to the 1st class, the trained sub-neural network must be different between the 1st and the 5th classes. Therefore, four sub-neural networks are able to correctly distinguish five classes from the former classified data. Across all 34 eigenvalues, the color feature provided better differentiation between the citrus HLB and non-citrus HLB, so only color features were applied in the four sub-networks. The testing results are showed in Figure 5 where the green dots represent the actual output and the blue lines represent the expected output of 1, 2, 3, 4 and 5. Specifically, the area of 1-50 represents the class of uniform yellowing, the area of 51-100 represents the floral symptom, 101-150 represents the healthy (normal mature), 151-200 represents the blocky mottled, and 201-250 stands for the yellowing of non-HLB. The recognition time was about 10 s. From Figure 5 and Table 3, the classification accuracy was approximately 90%.

HLB detection accuracy of the two-stage BPNN within the yellowing of non-HLB class reached over 90% (Table 3). The combination of gray features and color features performed better than color co-occurrence feature extraction method[2]. However, the number of HLB-positive and HLB-negative symptoms involved in this study was limited. There are other symptoms that

need to be considered, such as nutrient-deficiency in HLB-positive leaves. For other symptoms, the proposed feature extraction method in this study might not be appropriate, as the image of a nutrient-deficient HLB-positive leaf is quite similar to that of a nutrient-deficient HLB-negative leaf. A detection method based only on visual image data might not provide early detection of HLB because typical, early HLB symptoms are not obvious. Additionally, different citrus varieties varying in appearance means that the texture and color features that need to be extracted are likely to be different. Multispectral or hyperspectral data, combined with visual image data, is suggested for dealing with more different varieties of HLB citrus leaf symptoms. Consequently, a detection method based only on visual image data, using the same feature extraction method applied to all leaf samples, is not feasible.

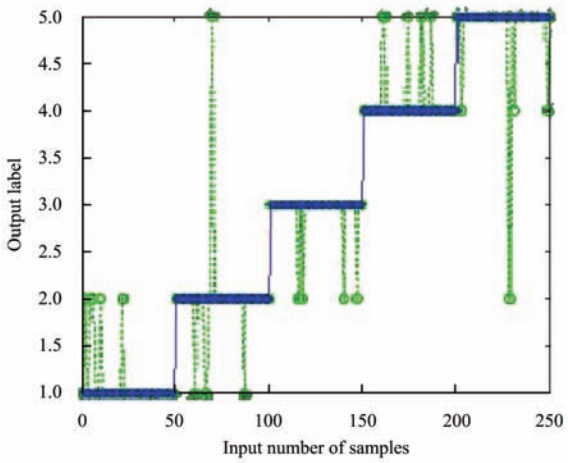



Figure 5 Classification results of five neural networks

**Table 3 The final classification accuracy**

unit: %

| | Uniform yellowing | Floral | Normal mature | Blocky mottled | Yellowing of non-HLB |
|---|---|---|---|---|---|
| Uniform yellowing | 90 | 10 | 0 | 0 | 0 |
| Floral | 8 | 90 | 0 | 0 | 2 |
| Normal mature | 0 | 8 | 92 | 0 | 0 |
| Blocky mottled | 0 | 0 | 0 | 90 | 10 |
| Yellowing of non-HLB | 0 | 2 | 0 | 6 | 92 |

None of the existing methods, even time-consuming, labor intensive, and relatively expensive laboratory-based detection methods such as qrt-PCR (quantitative real-time polymerase chain reaction) analysis, is 100% accurate. In this study, an easy method that can be easily extended as a real-time, portable, in-field diagnostic tool with an acceptable accuracy was presented. However, since three HLB-positive symptoms typically exist in real citrus orchards, future work will focus on additional disease symptom classes as well as additional citrus varieties, and will continue to explore more effective feature extraction methods to discover the heterogeneity of HLB.

## 4 Conclusions

In this study, an HLB detection method based on image feature extraction and two-stage BPNN was developed. Thirty-four features of sand sugar citrus were extracted to express the heterogeneity of HLB. Three typical symptoms of HLB and five classes of HLB-negative symptoms, including healthy leaves, were modeled through a two-stage BPNN. The results showed that the discrimination accuracy of HLB leaves from the non-HLB leaves with similar symptoms was acceptable.

## Acknowledgement

The authors thank for the supporting of National Natural Science Foundation of China: The research of citrus Huanglongbing in-field detection based on low-altitude multi-sensor fusion (Grant No.61675003), and the National Key Research and Development Plan: High efficient ground and aerial spraying technology and intelligent equipment (Grant No. 2016YFD0200700).

## [References]

[1] Kumar A, Lee W S, Ehsani R, Albrigo L G, Yang C, Mangan R L. Citrus greening disease detection using airborne multispectral and hyperspectral imaging. International Conference on Precision Agriculture, Denver, Colorado USA. 2010.

[2] Fan G, Liu B, Wu R, Li T, Cai Z, Ke C. Thirty years of research on citrus Huanglongbing in China. Fujian Journal of Agricultural Sciences, 2009; 24(2): 183–190. (in Chinese with English abstract)

[3] Durborow S. An analysis of the potential economic impact of Huanglongbing on the California citrus industry. Southern Agricultural Economics Association Annual Meeting, Orlando, FL, 2013-2-3.

[4] Gao Y, Lu Z, Liu Z, Zhong B. Research progress on diagnostic methods of citrus Huanglongbing. Journal of Gannan Normal University, 2013; 3: 37–40. (in Chinese

with English abstract)

[5] Chen Z, Li D. Preliminary studies on methods for rapid diagnosis of citrus yellow shoot disease (CYS). Journal of Zhejiang Agricultural University, 1987; 13(4): 348–354. (in Chinese)

[6] Zhang W. Study on PCR detection of citrus Huanglongbing bacterium. Biological Disaster Science, 2012; 35(2): 164–168.

[7] Pereira F, Milori D, Pereira-Filho E, Venâncio A, Russo M , Cardinali M, et al. Laser-induced fluorescence imaging method to monitor citrus greening disease. Computers and Electronics in Agriculture, 2011; 79(1): 90–93.

[8] Deng X, Li Z, Deng X, Hong T. Citrus disease recognition based on weighted scalable vocabulary tree. Precision Agriculture, 2014; 15(3): 321–330.

[9] Kim D G, Burks T F, Schumann A W, Zekri M, Zhao X H, Qin J. Detection of citrus greening using microscopic imaging. Agricultural Engineering International: the CIGR Journal, 2009; Manuscript 1194. Vol. XI. June.

[10] Pourreza A, Lee W S, Raveh E, Ehsani R, Etxeberria E. Citrus Huanglongbing detection using narrow-band imaging and polarized illumination. Transactions of the ASABE, 2014; 57(1): 259–272.

[11] Mei H, Deng X, Hong T, Luo X, Deng X. Early detection and grading of citrus Huanglongbing using hyperspectral imaging technique. Transactions of the CSAE, 2014; 30(9): 140–147. (in Chinese with English abstract)

[12] Deng X, Zheng J, Mei H, Li Z, Deng X, Hong T. Identification and classification of citrus Huanglongbing disease based on hyperspectral imaging. Journal of Northwest A&F University (Nat. Sci. Ed.), 2013; 41(7): 99–105. (in Chinese with English abstract)

[13] Deng X, Kong C, Wu W, Mei H, Li Z, Deng X, et al. Detection of citrus Huanglongbing based on principal component analysis and back propagation neural network. Acta Photonica Sinica, 2014; 43(4): 1–7. (in Chinese with English abstract)

[14] Li X, Lee W S, Li M, Ehsani R, Mishra A R, Yang C, et al. Comparison of different detection methods for citrus greening disease based on airborne multispectral and hyperspectral imagery. ASABE Paper. 2011; No.1110570. St. Joseph, Mich.: ASABE.

[15] Li X, Lee W S, Li M, Ehsani R, Mishra A R, Yang C, et al. Spectral difference analysis and airborne imaging classification for citrus greening infected trees. Computers and Electronics in Agriculture, 2012; 83(4): 32–46.

[16] Li H, Lee W S, Wang K, Ehsani R, Yang C. Extended spectral angle mapping (ESAM) for citrus greening disease detection using airborne hyperspectral imaging. Precision Agric, 2014; 15(2): 162–183.

[17] Li H, Lee W S, Wang R, Ehsani R, Yang C. Spectral angle mapper (SAM) based citrus greening disease detection using airborne hyperspectral imaging. Proc. 11th Intl Conf. on Precision Agriculture. Monticello, Ill.: International Society of Precision Agriculture, 2012.

[18] Lu Y, Guan H, Zhao B, Yang L. Study on the method of image pre-processing and feature extraction for rice diseases. Journal of Agricultural Mechanization Research, 2011; 33(8): 27–30. (in Chinese with English abstract)

[19] Wu S. Based on sobel edge detection operator of matlab implementation. Computer Knowledge and Technology, 2010; 16(19): 5314–5315.

[20] Guo D, Song Z. A Study on texture image classifying based on gray-level co-occurrence matrix. Forestry Machinery & Woodworking Equipment, 2005; 33(7): 21–23. (in Chinese with English abstract)

[21] Zhang J, Wang S, Dong X, Cheng P. A Study on Method of Extract of Texture Characteristic Value in Image Processing for Plant Disease of Greenhouse. Journal of Shenyang Agricultural University, 2006; 37(3): 282–285. (in Chinese with English abstract)

[22] Yang J, Liu C. Study of color space and its conversions in digital image processing. Journal of Shangqiu Vocational and Technical College, 2009; 8(2): 25–31. (in Chinese with English abstract)