Classification of grapefruit peel diseases using color texture feature analysis

Dae Gwan Kim, Thomas F. Burks, Jianwei Qin, Duke M. Bulanon

(*Department of Agricultural and Biological Engineering, University of Florida, Gainesville, FL* 32611-0570, *USA*)

Abstract: Technologies that can efficiently identify citrus diseases would assure fruit quality and safety and minimize losses for citrus industry. This research was aimed to investigate the potential of using color texture features for detecting citrus peel diseases. A color imaging system was developed to acquire RGB images from grapefruits with normal and five common diseased peel conditions (i.e., canker, copper burn, greasy spot, melanose, and wind scar). A total of 39 image texture features were determined from the transformed hue (H), saturation (S), and intensity (I) region-of-interest images using the color co-occurrence method for each fruit sample. Algorithms for selecting useful texture features were developed based on a stepwise discriminant analysis, and 14, 9, and 11 texture features were selected for three color combinations of HSI, HS, and I, respectively. Classification models were constructed using the reduced texture feature sets through a discriminant function based on a measure of the generalized squared distance. The model using 14 selected HSI texture features achieved the best classification accuracy (96.7%), which suggested that it would be best to use a reduced hue, saturation and intensity texture feature set to differentiate citrus peel diseases. Average classification accuracy and standard deviation were 96.0% and 2.3%, respectively, for a stability test of the classification model, indicating that the model is robust for classifying new fruit samples according to their peel conditions. This research demonstrated that color imaging and texture feature analysis could be used for classifying citrus peel diseases under the controlled laboratory lighting conditions.

Keywords: Citrus, disease detection, machine vision, color co-occurrence method, texture features, discriminant analysis
DOI: 10.3965/j.issn.1934-6344.2009.03.041-050

Citation: Dae Gwan Kim, Thomas F. Burks, Jianwei Qin, Duke M. Bulanon. Classification of grapefruit peel diseases using color texture feature analysis. Int J Agric & Biol Eng, 2009; 2(3): 41－50.

1 Introduction

During the recent past, citrus canker has become serious threats to citrus in Florida. These diseases can result in tree decline, death, yield loss and lost marketability. An automated detection system may help in citrus canker prevention for rapid and objective detection the diseases and, thus reduce the serious loss to the Florida citrus industry.

The automation of disease detection is needed for two major reasons. First, commercial citrus is produced on large farms. This needs significant demands for labor to manage diseases and pests. Consequently, labor saving tools are necessary to effective scout for and manage diseases and pests. These have stimulated development of automated scouts which can operate effectively for long periods of time. These systems can be more cost effective and accurate than human scouts. In addition, automated disease detection can help farmers reduce cost. The second is the continual inflow of non-native disease

Received date: 2009-04-30 **Accepted date:** 2009-09-03
Biographies: Dae Gwan Kim, graduate student, Email: dgkim030 @ufl.edu; **Thomas F. Burks**, Ph.D, professor, Department of Agricultural and Biological Engineering, University of Florida, Gainesville, FL 32611-0570, USA. 233 Frazier Rogers Hall, P.O. Box 110570, Email: tburks@ufl.edu; **Jianwei Qin**, Ph.D, Post-Doctoral Research Associate, email: qinj@ufl.edu; **Duke M. Bulanon**, Ph.D, Post-Doctoral Research Associate, Email: bulanon@ufl.edu

and insects. Non-native species can wreak havoc on local environmental and agricultural resources. Citrus canker entered into Florida using these international sources. Scouting for infected plants and insects by human inspectors has limited usefulness. Therefore, automated disease detection has several good points, including lowering production cost and improving scouting efficiency. Citrus trees can exhibit a host of symptoms reflecting various disorders that can adversely impact their health, vigor, and productivity to varying degrees. In some cases, disease control actions or remedial measures can be undertaken if the symptoms are identified early. Additional opportunities for disease control exist when precision agriculture techniques are involved, which could use early detection along with a global positioning system to map diseases in the grove for future control actions. This study explored machine vision based techniques that can visually differentiate common citrus peel diseases using individual fruit color texture features. Citrus samples were collected in the field and evaluated under laboratory conditions. Future studies will expand the technologies to in-field inspections.

In the past decade, various researchers have used image processing and pattern recognition techniques in agricultural applications, such as detection of weeds in the field, and sorting of fruits and vegetables. The underlying approach for all of these techniques is the same. First, images are acquired from the environment using analog, digital, or video cameras. Then, image processing techniques are applied to extract useful features that are necessary for further analysis of the images. Afterwards, discriminant techniques, such as parametric or non-parametric statistical classifiers and neural networks, are employed to classify the images. The selection of the image processing techniques and the classification strategies are important for the successful implementation of any machine vision system. Object shape matching functions, color-based classifiers, reflectance-based classifiers, and texture-based classifiers are some of the common methods that have been tried in the past.

A number of techniques have been studied to detect defects and diseases related to citrus. Gaffney (1973)[1] obtained reflectance spectra of citrus fruit and some surface defects. Edwards and Sweet (1986)[2] developed a method to assess damages due to citrus blight disease on citrus plants using reflectance spectra of the entire tree. Miller and Drouillard (2001)[3] collected data from Florida grapefruit, orange, and tangerine varieties using a color vision system. They used various neural network classification strategies to detect blemish-related features for the citrus fruit. Aleixos et al. (2002)[4] developed a multispectral camera system that could acquire visible and near infrared images from the same scene, and used it on a real-time system for detecting defects on citrus surface. Blasco et al. (2007)[5] reported the application of near-infrared, ultraviolet and fluorescence computer vision systems to identify the common defects of citrus fruit. They proposed a fruit sorting algorithm that combines the different spectral information to classify fruit according to the type of defect. Their results showed that non-visible information can improve the identification of some defects. Most recently, Qin et al. (2008)[6] developed an approach for citrus canker detection using hyperspectral reflectance imaging and PCA-based image classification method. Their results demonstrated that hyperspectral imaging technique could be used for discriminating citrus canker from other confounding diseases.

This research was aimed to develop a method to detect citrus peel diseases using color texture features. The use of color texture features in classical gray image texture analysis was first reported by Shearer (1986)[7]. Shearer and Holmes (1990)[8] reported a study for classifying different types of nursery stock by the color co-occurrence method (CCM). This method had the ability to discriminate between multiple canopy species and was insensitive to leaf scale and orientation. The use of color features in the visible light spectrum provided additional image characteristic features over traditional gray-scale texture representation. The textural methods employed were statistical-based algorithms that measured image features, such as smoothness, coarseness, graininess, and so on. The CCM method involved three major mathematical processes in

the following, and a complete discussion of the color co-occurrence method could be found in Shearer and Holmes (1990)[8].

1) Transformation of a red, green, blue (RGB) color representation of an image to an equivalent hue, saturation, and intensity (HSI) color representation;

2) Generation of color co-occurrence matrices from the HSI pixel maps. Each HSI matrix is used to generate a spatial gray-level dependence matrix (SGDM) providing three SGDM's.

3) Calculation of texture features from the three SGDM's.

Burks et al. (2000)[9] developed a method for weed species classification using color texture features and discriminant analysis. In their study, CCM texture feature data models for six classes of ground cover (giant foxtails, crabgrass, velvet leaf, lambs quarter, ivy leaf morning glory, and soil) were developed and stepwise discriminant analysis techniques were utilized to identify combinations of CCM texture feature variables, which have the highest classification accuracy with the least number of texture variables. A discriminant classifier was trained to identify weeds using the models generated. Classification tests were conducted with each model to determine their potential for classifying weed species. Pydipati et al. (2006)[10] utilized the color co-occurrence method to extract various textural features from the color RGB images of citrus leaves. The CCM texture statistics were used to identify diseased and normal citrus leaves using discriminant analysis.

The overall objective of this research was to develop a machine vision based method for detecting various diseases on citrus peel using color texture features under a controlled lighting condition. Specific objectives implemented to accomplish the overall objective were to: use a color imaging system to collect RGB images from grapefruits with normal and five diseased peel conditions (i.e., canker, copper burn, greasy spot, melanose, and wind scar); determine image texture features based on the color co-occurrence method (CCM); and develop algorithms for selecting useful texture features and classifying the citrus peel conditions based on the reduced texture feature sets.

2 Materials and methods

2.1 Citrus samples

Grapefruit is one of the citrus varieties that are susceptible to common peel diseases. Ruby Red grapefruits were used in this study. Fruit samples were handpicked from a grapefruit grove near Punta Gorda, Florida, during the harvest season in spring of 2007. The grapefruits with normal and five diseased peel conditions (i.e., canker, copper burn, greasy spot, melanose, and wind scar) were collected. Representative images for each peel condition are shown in Figure 1. Thirty samples for each condition were selected, hence a total of 180 grapefruits were tested in this study. All the grapefruits were washed and treated with chlorine and sodium o-phenylphenate (SOPP) at the Indian River Research and Education Center of University of Florida in Fort Pierce, Florida. The samples were then stored in an environmental control chamber maintained at 4℃ and they were removed from cold storage about 2 hours before imaging to allow them to reach room temperature.

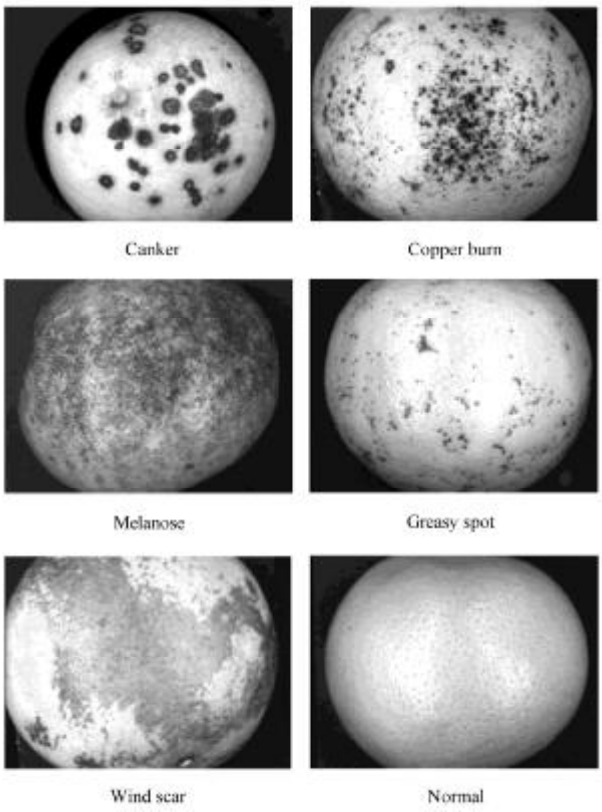

Figure 1 Normal and diseased citrus peel conditions

2.2 Color image acquisition

A color image acquisition system was assembled for acquiring RGB images from citrus samples, and it is shown in Figure 2. The imaging system consisted of two 13 W high frequency sealed fluorescent lights (SL Series, StockerYale, Salem, NH, USA), a zoom lens (Zoom 7000, Navitar, Rochester, NY, USA), a 3-CCD RGB color camera (CV-M90, JAI, San Jose, CA, UDA), a 24-bit color frame grabber board with 480×640 pixel resolution (PC-RGB, Coreco Imaging, St. Laurent, Quebec, CA), and a computer installed with an image capture software (MVTools, Coreco Imaging). The setup of the lighting system was designed to minimize specular reflectance and shadow and to maximize the contrast of the images. The height of the camera and its focus were adjusted to contain the image of the whole fruit, with an approximate 100 mm×100 mm field of view. Automatic white balance calibration was conducted using a calibrated white balance card before acquiring images from fruit samples. The digital color images were saved in uncompressed BMP format.

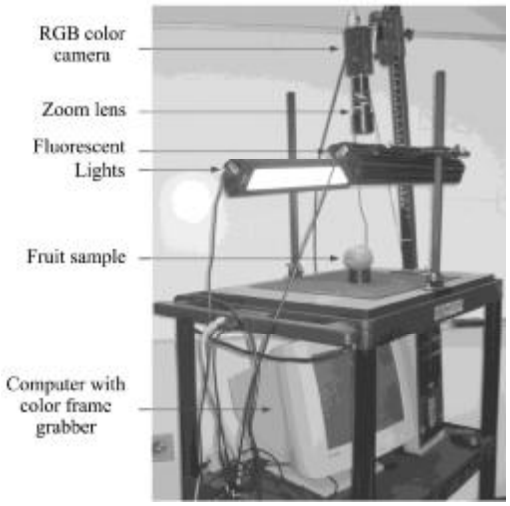

Figure 2 Color image system for acquiring RGB images from citrus samples

2.3 Data analysis for color images

The data analysis methods for analyzing the color images of the fruit samples based on the color co-occurrence method (CCM) are illustrated in the flow chart shown in Figure 3, which involve the procedures for selection of region of interest (ROI), transformation from RGB format to HSI format, generation of spatial gray-level dependence matrices (SGDM's), calculation of texture features, selection of useful texture features, and discriminant analysis for disease classification. All image processing and data analysis procedures were executed using programs developed in Matlab 7.0 (MathWorks, Natick, MA, USA) and SAS 9.1 (SAS Institute Inc., Cary, NC, USA)[11]. Detailed methods and procedures for each step are described in the following sections.

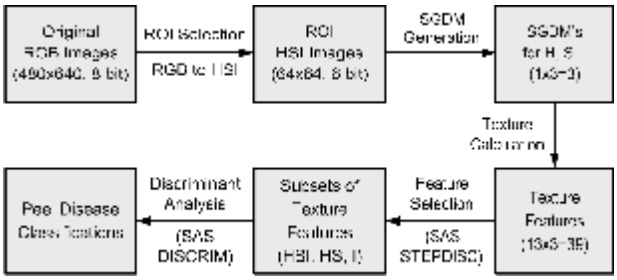

Figure 3 Procedures for color image analysis

2.4 ROI selection and color space conversion

ROI images were first extracted from the original RGB color images with the dimension of 480×640 pixels, generating small images covering the interested areas (i.e., normal peels or various diseases) on the fruit surface. The ROI selection was started manually by determining a point on the original image, and then was finished by a Matlab program for extracting a square portion with the dimension of 64×64 pixels centered on the determined point. This approach obtains the useful image data and significantly reduces the computational burden for the following data analysis procedures. Representative ROI images for each fruit peel condition used in this study are shown in Figure 4. Figure 5 shows 15 representative images about canker condition.

For reducing the computational burden with minimal loss of texture feature quality, the ROI images were then converted from the original eight bit per channel red, green, blue (RGB) color representation to a six bit per channel hue, saturation, and intensity (HSI) color representation to facilitate the SGDM calculation. Intensity is calculated using the mean value of the three RGB values. The hue and saturation values are

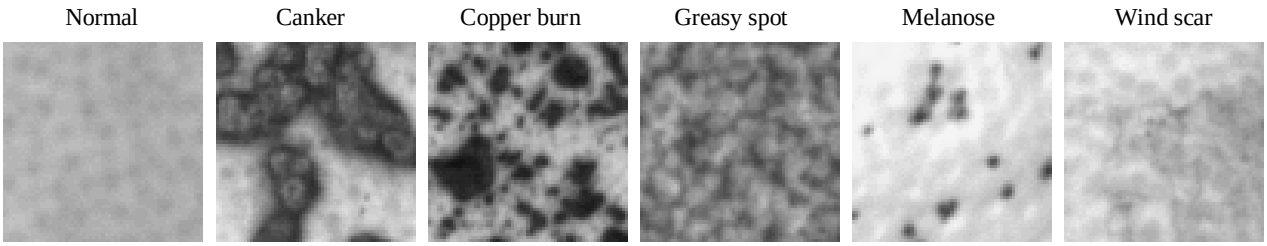

Figure 4 Typical ROI images for normal and diseased citrus peel conditions

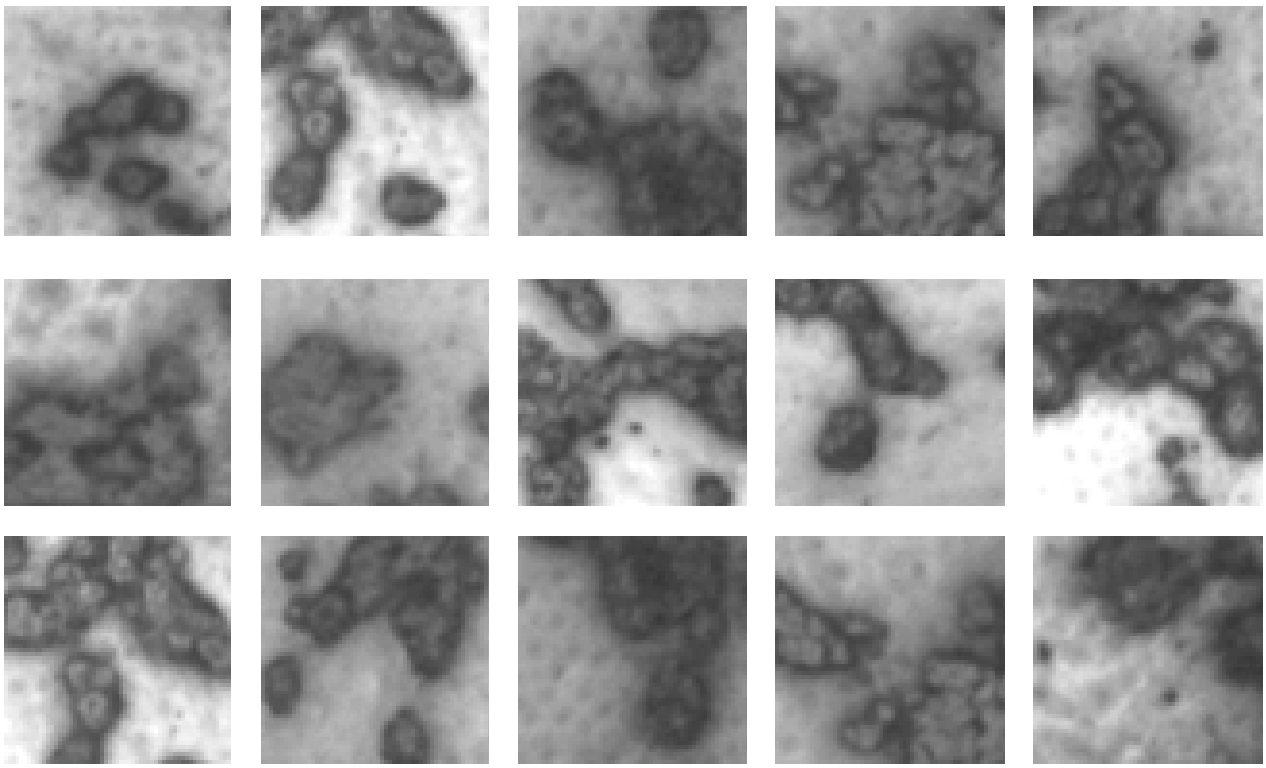

Figure 5 15 ROI images for citrus canker condition

determined using a geometrical transformation of the International Commission on Illumination's (CIE) chromaticity diagram (Ohta, 1985)[12]. In this process, the CIE chromaticity diagram represents a two-dimensional hue and saturation space (Wyszecki and Stiles, 1992)[13]. The RGB values determine the chromaticity coordinates on the hue and saturation space, which are then used to geometrically calculate the value of hue and saturation.

2.5 SGDM generation

Before applying CCM method to input images, the original image consisting of a red, green, and blue (RGB) color space are converted to a HSI color space. HSI space is distributed into hue, saturation, and intensity components. Most of image processing engines and methods are based on HSI color space system. This color system has strong tolerance for a change of a light on an image or a reflection. This characteristic of HSI can help image processing be less sensitive to illumination of surroundings. The color co-occurrence texture analysis method was developed through the use of the spatial gray-level dependence matrices (SGDM's). The SGDM's were generated for each color pixel map of the ROI HSI images, one each for hue, saturation and intensity. These matrices measure the probability that a pixel at one particular gray-level will occur at a distinct distance and orientation from any pixel given that pixel has a second particular gray-level (Shearer and Holmes, 1990)[8]. The SGDM is represented by the function $P(i,$

j, *d*, θ) where i represents the gray-level of the pixel at (x_1, y_1) in the image, and *j* represents the gray-level of the pixel at (x_2, y_2) located at a distance d and an orientation angle θ from (x_1, y_1) (Shearer, 1986)[7]. The matrix is constructed by counting the number of pixel pairs of (x_1, y_1) and (x_2, y_2) with the grey value *i* and *j* at distance *d* and direction θ. An example for the nearest neighbor mask is illustrated in Figure 6, where the reference pixel is shown as an asterisk. All eight neighbors are one pixel distance from the reference pixel '*' and are numbered in a clockwise direction from one to eight. The neighbors at positions one and five are both considered to be at an orientation angle equal to 0°, while positions eight and four are considered to be at an angle of 45°. It was determined from the preliminary test that the calculation would use a 0° orientation angle and an offset distance of one pixel. The offset represents the coarseness of the texture evaluation, where the smaller the offset, the finer the texture measured. Thus the one pixel offset is the finest texture measure.

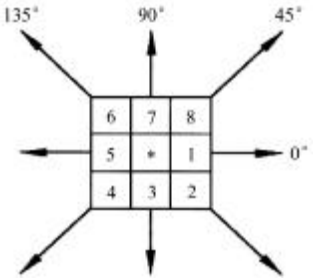

Figure 6 Nearest neighbor mask for calculating spatial gray-level dependence matrices (SGDM's)

2.6 Texture feature calculation

The SGDM's generated for hue, saturation and intensity were then used to calculate the texture features. Shearer and Holmes (1990)[8] reported a reduction for the 16 gray scale texture features through elimination of redundant variables, resulting in 11 texture features. Donohue et al. (2001)[14] added two more texture features (i.e., image contrast and modus) to those used by Shearer and Holmes (1990)[8]. In this study, the combined 13 texture features proposed by Shearer and Holmes (1990)[8] and Donohue et al. (2001)[14] were used for citrus peel disease classification, and they included (1) uniformity, (2) mean intensity, (3) variance, (4) correlation, (5) product moment, (6) inverse difference, (7) entropy, (8) sum entropy, (9) difference entropy, (10) information correlation #1, (11) information correlation #2, (12) contrast, and (13) modus. The equations for calculating the 13 texture features can be found in Pydipati et al. (2006)[10].

The calculations were performed for each of the three SGDM's, producing 13 texture features for each HSI component and thereby a total of 39 texture statistics. The texture features were identified by a coded variable name where the first letter represents whether it is a hue (H), saturation (S) or intensity (I) feature and the number following represents one of the thirteen texture features described above. Intensity texture feature equations are presented in Table 1. As an example, the feature (I_7) is a measure of the entropy in the intensity CCM matrix, which represents the amount of order in an image and is calculated by equation 1.

$$I_7 = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} p(i.j) \ln\ p(i,j) \quad (1)$$

The $p(I,j)$ matrix represents the normalized intensity co-occurrence matrix and N_g represents the total number of intensity levels. The equation for normalizing the co-occurrence matrix is given in equation 2, where $P(i,j,1,0)$ is the intensity co-occurrence matrix.

$$p(i,j) = \frac{P(i,j,1,0)}{\sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} P(i,j,1,0)} \quad (2)$$

A physical representation of entropy (uncertainty) may be visualized by comparing a checkerboard-like image to an image where one half is black and the other half is white. The latter image is highly ordered having all pixels of the same intensity segregated into two distinct pixels groups, which gives greater certainty of the pixel value of the adjacent pixels. The checkerboard image has a lower amount of order due to intermixing of black and white squares, which results in a greater level of uncertainty of neighboring pixel values. The lower order image would therefore have more uncertainty and thus a higher entropy measure.

Table 1 Intensity texture features

Feature	Description	Equation
I1	Uniformity (2^{nd} Moment)	$I_1 = \sum_{i=0}^{N_g-1}\sum_{j=0}^{N_g-1}[p(i,j)]^2$
I2	Mean intensity	$I_2 = \sum_{i=0}^{N_g-1} ip_x(i)$
I3	Variance	$I_3 = \sum_{i=0}^{N_g-1}(i-I_2)p_x(i)$
I4	Correlation	$I_4 = \frac{\sum_{i=0}^{N_g-1}\sum_{j=0}^{N_g-1} ijp(i,j) - I_2^{\,2}}{I_3}$
I5	Product moment	$I_5 = \sum_{i=0}^{N_g-1}\sum_{j=0}^{N_g-1}(i-I_2)(j-I_2)p(i,j)$
I6	Inverse difference	$I_6 = \sum_{i=0}^{N_g-1}\sum_{j=0}^{N_g-1}\frac{p(i,j)}{1+(i-j)^2}$
I7	Entropy	$I_7 = \sum_{i=0}^{N_g-1}\sum_{j=0}^{N_g-1}p(i,j)\ln p(i,j)$
I8	Sum entropy	$I_8 = \sum_{k=0}^{2(N_g-1)} p_{x+y}(k)\ln p_{x+y}(k)$
I9	Difference entropy	$I_9 = \sum_{k=0}^{N_g-1} p_{x-y}(k)\ln p_{x-y}(k)$
I10	Information correlation 1	$I_{10} = \frac{I_7 - HXY1}{HX}$
I11	Information correlation 2	$I_{11} = \left[1-e^{-2(HXY2-I_7)}\right]^{1/2}$
	HX	$HX = -\sum_{i=0}^{N_g-1} p_x(i)\ln p_x(i)$
	HXY1	$HXY1 = -\sum_{i=0}^{N_g-1}\sum_{j=0}^{N_g-1}p(i,j)\ln\left[p_x(i)p_x(j)\right]$
	HXY2	$HXY2 = -\sum_{i=0}^{N_g-1}\sum_{j=0}^{N_g-1}p_x(i)p_x(j)\ln\left[p_x(i)p_x(j)\right]$
I12	Contrast	$I_{12} = \sum_{\lvert i-j\rvert=0}^{N_g-1}(i-j)^2\sum_{i=0}^{N_g-1}\sum_{j=0}^{N_g-1}p(i,j)$
I13	Modus	$I_{13} = \sum_{i=0}^{N_g-1}\sum_{j=0}^{N_g-1}\max\left[p(i,j)\right]$

2.7 Texture feature selection

After the texture statistics were obtained for each image, feature selection was conducted to reduce the redundancy in the texture feature set. The SAS procedure STEPDISC can reduce the size of the variable set and find the variables that are important for discriminating samples in different classes, and it was used for the texture feature selection. In the STEPDISC procedure, forward selection, backward elimination, or stepwise selection are used for analyses (Klecka, 1980)[15]. The stepwise discriminant analysis begins with no variables in the classification model for this research. At each step of the process, the variables within and outside the model are evaluated. The variable within the model, at that particular step, which contributes least to the model as determined by the Wilks' Lambda method is removed from the model. Likewise, the variable outside the model that contributes most to the model and passes the test to be admitted is added. A test significant level of 0.0001 for the variables of SLS (test for variable to stay) and the SLE (test for variable to enter) in the STEPDISC procedure was chosen for the stepwise discrimination of the variable list (SAS, 2004)[11]. When no more steps can be taken, the number of variables in the model is reduced to its final form.

Burks et al. (2000)[9] had shown that classification performances were poor if only hue or saturation information was used in the classification models. Thus three color feature combinations including hue, saturation, and intensity (H, S, I), hue and saturation (H, S), and intensity (I) only were used to perform the texture feature selections. These three color combinations have demonstrated high classification accuracies in the applications for other plant discriminations[9,10].

2.8 Texture classification

The classification models were developed using the SAS procedure DISCRIM, which creates a discriminant function based on a measure of the generalized squared distance between a specific test image texture variable input set and the class texture variable means, with an additional criteria being the posterior probability of the classification groups (Rao, 1973)[16]. Each sample in the testing set was placed in the class for which it had the smallest generalized square distance between the test observation and the selected class, or the largest posterior probability of being in the selected class. The DISCRIM procedure utilized a likelihood ratio test for homogeneity of the within-group covariance matrices at a 0.1 test significance level.

The 30 samples from each peel condition were divided into two datasets consisting of 20 samples for training and 10 samples for testing. The samples were first arranged in ascending order for the time the images were acquired. The first two samples were selected for training and the third sample for testing. This approach

minimizes negative time dependent variability, and reduces potential for data selection bias between the training and test datasets. A training data set and a test data set were created for each of the subsets of the texture features selected by the stepwise discriminant analysis described above. The training sets were used to train the classification models and the testing sets were used to evaluate the accuracies of different classification models.

3 Results and discussion

3.1 Selection of texture features

The texture feature selection results are summarized in Table 1. Four classification models were developed using the selected texture feature sets from the three color combinations [(H, S, I), (H, S), and (I)]. The variables listed in the column of "Texture Feature Set" in Table 2 were generated by the SAS STEPDISC procedure, and they were arranged in the descending order of the importance for the classification models. The subscript numbers indicate the texture statistics as the following: (1) uniformity, (2) mean intensity, (3) variance, (4) correlation, (5) product moment, (6) inverse difference, (7) entropy, (8) sum entropy, (9) difference entropy, (10) information correlation #1, (11) information correlation #2, (12) contrast, and (13) modus. As an example, H_9 represents the difference entropy of hue, and it is selected as the most important texture feature for the first two classification models developed using two different color combinations [(H, S, I), and (H, S)].

Table 2 Texture features selected by stepwise discriminant analysis

Classification model	Color feature	Texture feature set
HSI_13	H, S, I	H_9, H_{10}, I_{12}, S_7, I_3, I_2, S_{12}, I_{11}, I_1, I_8, S_1, H_2, H_5
HS_9	H, S	H_9, H_{10}, S_7, H_5, H_{11}, S_{12}, S_{11}, H_7, H_{13}
I_11	I	I_2, I_3, I_5, I_{10}, I_6, I_{13}, I_8, I_9, I_1, I_{11}, I_7
HSI_39	H, S, I	All 39 texture features (H_1-H_{13}, S_1-S_{13}, I_1-I_{13})

The classification models were named using the color features involved in the texture feature selections followed by the total numbers of the selected texture features. For example, model HSI_13 consists of a reduced set of hue, saturation and intensity texture features, and there are 13 texture features in total that were used to construct the model. As shown in Table 2, significant eliminations of redundant texture features were accomplished through the stepwise discriminant analysis. Nine and 11 texture features were selected for model HS_9 and model I_11, respectively. The simplification of the texture features largely reduces the computation burden due to the redundant data, and it also helps improve the performance of classification models. In addition to the three models described above, a classification model that used all 39 HSI texture features was developed for the purpose of comparisons with other models. Thus there are four classification models that were used to differentiate the citrus peel diseases, and they were independently evaluated for classification performance.

3.2 Classification of citrus peel conditions

The SAS procedure DISCRIM was used to test the accuracies of the classification models. Table 3 summarizes the classification results for differentiating different citrus peel conditions using model HSI_13 listed in Table 2. As shown in Table 3, four peel conditions (normal, canker, copper burn, and wind scar) among the total six conditions tested in this study were perfectly classified into the appropriate categories. For the other two conditions (greasy spot and melanose), there was one misclassified sample for each case. One greasy spot sample was misclassified as copper burn, and one melanose sample was misclassified as wind scar. The classification accuracies for greasy spot and melanose were 90%. In general, there were only two samples that were misclassified in the 60 samples in the testing set, and the overall classification accuracy for the model HSI_13 was 96.7%.

Same procedures were applied for the other three classification models listed in Table 2, and the classification results, along with those from the model HSI_13, are summarized in Table 3. Using nine selected hue and saturation texture features, model HS_9 provided the classification accuracies of 90% for normal, canker, copper burn, greasy spot, and wind scar, and 70% accuracy for melanose. The average accuracy of the

model HS_9 was 86.7%. Model I_11 used 11 selected intensity texture features alone. Although it achieved two perfect classification results (100%) for copper burn and greasy spot, the performances for the other four conditions were poor, especially for melanose (70%) and wind scar (50%). The overall accuracy of the model I_11 was 81.7%, which is the worst among the four models tested in this study. When all 39 texture features were used by model HSI_39, the classification accuracy was achieved as 88.3%, which was higher than those of the models HS_9 and I_11, but lower than that of the model HSI_14 (96.7%).

Table 3 Classification results using model HSI_13 in Table 2

Actual peel condition	Classified peel condition						Accuracy/%
	Normal	Canker	Copper burn	Greasy spot	Melanose	Wind scar	
Normal	10	0	0	0	0	0	100
Canker	0	10	0	0	0	0	100
Copper burn	0	0	10	0	0	0	100
Greasy spot	0	0	1	9	0	0	90
Melanose	0	0	0	0	9	1	90
Wind scar	0	0	0	0	0	10	100
Total	10	10	11	9	9	11	96.7

Table 4 Classification results in percent correct for all models in Table 1

Peel condition	Classification model			
	HSI_13	HS_9	I_11	HSI_39
Normal	100	90	80	80
Canker	100	90	90	100
Copper burn	100	90	100	90
Greasy spot	90	90	100	90
Melanose	90	70	70	70
Wind scar	100	90	50	100
Overall accuracy/%	96.7	86.7	81.7	88.3

Texture feature selection is necessary for obtaining better classification accuracy, and this is confirmed by the fact that the model using 13 selected hue, saturation and intensity texture features (model HSI_13) outperformed other classification models that used intensity texture features, (model I_11), hue and saturation texture features (model HS_9) and all 39 HSI texture features (model HSI_39).

3.3 Stability test of the classification model

The classification results presented in section 3.2 were obtained using the testing samples in a fixed order (i.e., one from every three samples arranged in ascending order for the time the images were acquired). To test the stability of the classification model, 20 training samples and 10 testing samples were randomly chosen from the 30 samples for each peel condition, and they were used to train and test the model HSI_13, which gave the best classification performance, following the same procedures described earlier. Ten runs were repeated for the training and testing. The average value and standard deviation in Table 5 were 96.0% and 2.3%, respectively.

Table 5 Classification results for shuffle data models in percent correct

Number of random data	Canker	Copper	Greasy spot	Market	Specular melanose	Windscar	Total
1	100	90	100	100	100	100	98.3
2	100	100	100	80	90	100	95.0
3	100	100	100	100	90	100	98.3
4	80	100	100	90	90	100	93.3
5	100	100	90	100	90	90	95.0
6	100	100	100	100	100	100	100
7	100	90	100	100	100	90	96.7
8	100	100	100	90	100	80	95.0
9	100	90	100	90	90	100	95.0
10	80	100	100	100	90	90	93.3
Average accuracy/%							96.0

Note: There were 10 times shuffle models in percent correct.

4 Summary and conclusions

Significant eliminations of redundant texture features were accomplished through the stepwise discriminant analysis. 13, 9, and 11 texture features were selected for the color combinations of HSI, HS, and I, respectively. The simplification of the texture features largely reduces the computation burden, and it also helps improve the performance of classification models. The classification model using intensity texture features only gave the worst accuracy (81.7%), and the model using 13 selected HSI texture features achieved the best classification accuracy

(96.7%) among four classification models including the one using all 39 HSI texture features. The results suggested that it would be best to use a reduced hue, saturation and intensity texture feature set to differentiate different citrus peel conditions. A stability test for the classification model with the best performance was accomplished by 10 runs using randomly selected training and testing samples. Average classification accuracy and standard deviation in table 5 were 96.0% and 2.3%, respectively, indicating that the classification model is robust for classifying new fruit samples according to their peel conditions.

This research demonstrated that color imaging and texture feature analysis could be used for differentiating citrus peel diseases under the controlled laboratory lighting conditions. Future studies will explore the utility of these algorithms in outdoor conditions, and develop neural network methods such as self-organizing map (SOM) or support vector machines (SVM) for real-time application. The most significant challenge will be created by the inherent variability of color under natural lighting conditions. By eliminating intensity-based texture features, this variability can be significantly reduced. However, hue and saturation can be somewhat influenced by low lighting conditions. This may point to the need to use cameras with light availability color compensation, supplemental lighting, or night time applications where lighting levels can be controlled.

[References]

[1] Gaffney J J. Reflectance properties of citrus fruits. Trans. ASAE, 1973; 16(2): 310－314.

[2] Edwards J G, Sweet C H. Citrus blight assessment using a microcomputer: quantifying damage using an apple computer to solve reflectance spectra of entire trees. Florida Scientist, 1986; 49 (1): 48－53.

[3] Miller W M, Drouillard G. P. Multiple feature analysis for machine vision grading of Florida citrus. Applied Eng Agri, 2001; 17(5): 627－633.

[4] Aleixos N, Blasco J, Navarron F, Molto E. Multispectral inspection of citrus in real-time using machine vision and digital signal processors. Comput. Electron Agri, 2002; 33(2): 121－137.

[5] Blasco J, Aleixos N, Gomez J, Molto E. Citrus sorting by identification of the most common defects using multispectral computer vision. J Food Eng, 2007; 83(3): 384－393.

[6] Qin J, Burks T F, Kim M S, Chao K, Ritenour M A. Citrus canker detection using hyperspectral reflectance imaging and PCA-based image classification method. Sens. Instrum. Food Qual. Saf., 2008. doi: 10.1007/s11694-008-9043-3.

[7] Shearer S A. Plant identification using color co-occurrence matrices derived from digitized images. Ph.D. Thesis, The Ohio State University, Columbus, OH, USA. 1986.

[8] Shearer S A, Holmes R. G. Plant identification using color co-occurrence matrices. Trans ASAE, 1990; 33 (6): 2037－2044.

[9] Burks T F, Shearer S A, Payne F A. Classification of weed species using color texture features and discriminant analysis. Trans. ASAE, 2000; 43 (2): 441－448.

[10] Pydipati R, Burks T F, Lee W S. Identification of citrus disease using color texture features and discriminant analysis. Comput Electron Agri, 2006; 52(1-2): 49－59.

[11] SAS. 2004. SAS/STAT 9.1 User's Guide. SAS Institute Inc., Cary, NC, USA.

[12] Ohta Y. Knowledge-based interpretation of outdoor natural color scenes. Pitman Publishing Inc., Marshfield, MA, USA. 1985.

[13] Wyszecki G, Stiles W S. Color science: concepts and methods. In: Quantitative Data and Formulae, second ed. pp. 117–137. John Wiley and Sons, New York, NY, USA. 1992.

[14] Donohue K D, Huang L, Burks T F, Forsberg F, Piccoli C W. Tissue classification with generalized spectrum parameters. Ultrasound Med. Biol. 2001, 27 (11): 1505－1514.

[15] Klecka W R. Discriminant analysis. Sage university paper series on quantitative applications in the social sciences,Series No. 07-019, Beverly Hills: Sage publications. 1980.

[16] Rao C. R. Linear statistical inference and its application. John Wiley and Sons, New York, NY, USA. 1973.