Postharvest Biology and Technology 143 (2018) 119–128

Contents lists available at ScienceDirect

Postharvest Biology and Technology

journal homepage: www.elsevier.com/locate/postharvbio

Non-destructive recognition and classification of citrus fruit blemishes based on ant colony optimized spectral information

Yao Zhang[a,b,c], Won Suk Lee[b,*], Minzan Li[a], Lihua Zheng[a], Mark A. Ritenour[d]

[a] Key Laboratory of Modern Precision Agriculture Integration Research, MOE, China Agricultural University, Beijing, 100083, China
[b] Department of Agricultural & Biological Engineering, University of Florida, Gainesville, FL 32611, United States
[c] Institute of Remote Sensing and Geographic Information System, School of Earth and Space Sciences, Peking University, Beijing, 100871, China
[d] Indian River Research and Education Center, University of Florida, Fort Pierce, FL, 34945, United States

ARTICLE INFO

Keywords:
Spectroscopy
Feature selection
Surface blemish
Classification
Packinghouse

ABSTRACT

Fast and accurate assessment of citrus fruit blemishes is critical to improve fruit quality and company profitability of citrus packinghouses and juice processing plants. This study aimed to identify spectral signatures of healthy fruit, and fruit exhibiting symptoms or damage from Huanglongbing (HLB), melanose, oleocellosis (oil spot), wind scar, leafminer and rust mites. Fruit samples were classified using identified spectral information. The current work proposes a characteristic waveband selection method based on the combination of the ant colony optimization (ACO) algorithm and variable selection principles. Six characteristic wavebands for each type of citrus blemishes were determined. Two different classification methods were established by the acquired characteristic wavebands, including simple layer support vector machine (SVM) classification models and tree-type SVM models. After using the tree-type SVM models, classification accuracies of healthy, HLB, melanose, oil spot, wind scar, leafminer and rust mite categories were 98.4%, 90.8%, 95.2%, 92.0%, 90.8%, 95.2% and 96.8%, respectively. The proposed characteristic wavebands selection methods were therefore very effective in extracting features of citrus fruit with these blemishes and the tree-type SVM classification models made it possible to correctly classify the fruit with high detection accuracies and universality.

1. Introduction

The United States ranks fifth in citrus production in the world, with almost half of the production being from Florida (USDA, 2017). Citrus fruit in Florida for the fresh market is an important segment valued at over $150 million annually (USDA, 2017). The external appearance of fruit is one of the most important factors determining perceived quality and ultimate price of fresh fruit because consumers judge quality most often by sight. For this reason, USDA grade standards limit the amount of blemishes allowed on fruit for the fresh market (USDA AMS, 2018a). Usually, the cause of the fruit blemish (e.g., from wind scar, melanose, etc.) is not determined during grading operations. However, such information would be beneficial to assist producers and packers in adjusting practices such as fungicide or insecticide applications to mitigate such blemishes. Moreover, the majority of citrus grown in Florida is processed for juice and other byproducts. Citrus fruit are dumped into the juice extractor system and through a series of processes, such as washing, oil extraction of peel, juice extraction, pulp and seed removing and pasteurizing, the juice is produced and separated by quality. Disease affected fruit reduce pulp quality and safety as well as the major flavor chemicals, volatile (aroma), non-volatile (taste) and mouth feel of the juice (Bai et al., 2016; Baldwin et al., 2010; Dagulo et al., 2010; Raithore et al., 2015; Zhao et al., 2015). However, some external blemishes on the fruit are not associated with decay and do not impact internal fruit quality or safety, and therefore could be used for the juice products (Davies and Albrigo, 1999). Rapid and accurate identification and classification of blemishes of fruit in a large quantity hence become a crucially important step in the citrus packinghouses and juice extraction plants, and would assure the competitiveness and profitability of the citrus industry.

Detection of fruit blemishes is an important first step in fruit diagnoses and classification. Blasco et al. (2007) proposed a region-growing algorithm using color information to detect the most common external blemishes such as green mold, scarring and thrips. Morphological parameters and decision algorithms were then added, yielding gradual improvement of the results; successful identification of 11 types of citrus blemishes reached 86% (Blasco et al., 2009). López-García et al. (2010) used a method that combined color and texture information in a principal component analysis (PCA) model to detect peel blemishes in four citrus cultivars. The accuracy of the detection of individual

* Corresponding author.
E-mail address: wslee@ufl.edu (W.S. Lee).

https://doi.org/10.1016/j.postharvbio.2018.05.004
Received 11 December 2017; Received in revised form 9 May 2018; Accepted 9 May 2018

blemishes was 91.5%, while the classification ratio of damaged/sound samples was 94.2%. Kim et al. (2009) and Zhao et al. (2009) proposed an algorithm based on a stepwise discriminant analysis to select representative texture features. Three reduced feature models including hue, saturation, and intensity (HSI), hue and saturation (HS), and intensity (I) only were used to classify the different citrus peel conditions. Lu et al. (2015) introduced a segmentation method of circularity threshold on the chromatic aberration map of green blue (GB) components to separate blemishes from healthy regions on citrus fruit surfaces. Therefore, while many studies have focused on color, texture and morphological information to detect blemishes on citrus fruit, to our knowledge, even for the same type of fruit blemish, the symptoms may be varying in color, texture and morphology. Furthermore, different types of blemishes may appear similar (Hawkins et al., 2010). Thus, if we want to detect and classify multiple types of blemishes on citrus fruit, it is essential to find out more targeted information about each type of blemish.

Spectral analysis studies the interaction of a material with electromagnetic energy to determine its unique spectral properties (Pourreza et al., 2015). Blasco et al. (2007) classified citrus blemishes using color, near-infrared (NIR), ultraviolet (UV) and ultraviolet fluorescence (UVFL) information. Their experiments showed that some citrus blemishes could only be detected in a certain region of wavebands. Sankaran et al. (2013) used visible-near infrared and thermal imaging to detect Huanglongbing (HLB) disease in citrus trees. They used average reflectance values from the visible, NIR, and thermal bands as inputs to classify the HLB diseased and healthy trees. Bulanon et al. (2013) identified four key wavelengths that could discriminate citrus black spot (CBS) using correlation analysis based on the wavelength ratio and wavelength difference. The pattern recognition approaches had an overall accuracy of 92.0%. Kim et al. (2014) used spectral angle mapper (SAM) and spectral information divergence (SID) hyperspectral analysis approaches to classify fruit samples into two classes: CBS or non-CBS. The classification accuracy for CBS with the SAM approach was 97.9%, and 97.1% with SID. Qin et al. (2009) took advantage of a SID classification method, which was based on quantifying the spectral similarities by using a predetermined canker reference spectrum to differentiate canker from normal fruit peels and other citrus surface conditions. The overall classification accuracy was 96.2%. Later, Qin et al. (2012) developed a method with only two wavebands of 730 and 830 nm for real-time citrus canker detection which could be applied in an on-line commercial fruit sorting machine. The overall classification accuracy was 95.3%. Niphadkar et al. (2013) determined the detectable size limit for cankerous lesions using hyperspectral imaging approaches. Wetterich et al. (2016) combined fluorescence imaging spectroscopy (FIS) and a machine learning technique to discriminate citrus canker, HLB and other ordinary citrus blemishes that presented similar symptoms. The classification results were 97.8% when discriminating citrus canker from citrus scab, and 95.0% when discriminating HLB from zinc deficiency. These studies reveal that spectral technology is an important approach and has an outstanding performance in citrus blemish detection. However, for blemished fruit detection, only some specific blemishes with obvious unique symptoms were investigated. And a universal method which could be applied in extracting the features of various types of citrus fruit surface blemishes has not been investigated. Therefore, a suitable approach with strong universality for exploring characteristic wavelengths of healthy fruit and typical blemishes that are mostly present at citrus orchards is needed. Besides that, instead of trying to detect only one specific blemish among different conditions and symptoms, classification methods which could identify the types of citrus blemishes simultaneously are also required.

For feature selection, swarm intelligence optimization algorithm emerges with a simulation of simple organism's group behavior. Based on the studies about ant behavior, biologists found that individuals demonstrate limitations and randomness in activities, but the whole swarm shows a high degree of self-organization. This phenomenon has aroused great interest among scientists, and a new bionic algorithm simulating ant foraging behavior, called ant colony optimization (ACO), was introduced in the early 1990s (Colorni et al., 1991). ACO is a development of artificial intelligence and swarm intelligence with the signatures of distributed computing, positive information feedback and heuristic search. It has been successfully applied in many areas such as communication, network route, traffic, and so on (Bonabeau et al., 2000; Dorigo et al., 1996; Shen et al., 2005). At present, some scholars have used the ACO algorithm to conduct feature selection in numerous areas. Shen et al. (2005) proposed a modified ACO algorithm which used a binary path selection mechanism to choose variables in quantitative structure-activity relationship (QSAR) modeling. Allegrini and Olivieri, (2011) employed the concept of cooperative pheromone accumulation, which is typical of ACO selection methods, and optimizes Partial Least Squares (PLS) models using a pre-defined number of variables, employing a Monte Carlo approach to discard irrelevant sensors. Guo et al. (2014) employed ACO-PLS to explore optimally the efficient wavelength from the NIR spectroscopy of apple to develop models for predicting the soluble solid content (SSC) based on the features of heuristic global search and the random selection mechanism of Monte Carlo roulette. Hou et al. (2016) applied the ant colony clustering algorithm (ACCA) to solve the fuzziness of the multispectral image for grapevine leafroll disease (GLD) infected grapevines and successfully identified GLD from healthy ones. Moreover, some scholars compared the performance of ACO and other heuristic algorithms in feature selection. Ke et al., (2008) proposed an ACO-based algorithm and compared its performance of feature selection in rough set theory with the simulated annealing-, genetic algorithm-, and Tabu search-based algorithms. The results showed the ACO-based algorithm achieved better performance according to both the classification results and the number of features. Santana et al. (2010) found ACO performed better than genetic algorithm-based feature selection method for ensemble classifiers when the number of individual classifiers was small. Agrawal and Kaur (2018) compared ACO and Hybrid Particle Swarm Optimization in test case selection. The results indicated ACO outperforms Hybrid Particle Swarm Optimization in the calculating efficiency. Throughout these studies, ACO performs more flexible and efficient. It is especially suitable for the relatively small-scale problems (Xue et al., 2016).

While in the process of feature selection using ACO, most studies adopted Monte Carlo roulette to select the next feature, which is random and lacks connection with the variable information. In this approach, the selected features mostly depended on the performance of ACO. However, for variable selection, there are two factors which are quite important. One is discrepancies with the selected wavebands; and another is the representative ability of the whole spectral information. Based on these criteria, this study presented a new, modified ACO algorithm combined with correlation relationship between variables.

Based on these previous studies, the objectives of this study were the following:

1 To find a universal approach to extract distinctive spectral information of different citrus fruit surface blemishes. Successive projection algorithm (SPA) and an innovative wavebands selection method that takes comprehensive consideration of ACO and feature selection criterion were tested in this study.
2 To establish classification methods, using the selected characteristic wavebands, to classify different types of citrus surface blemishes and disease symptoms, including HLB, melanose oil spot, wind scar, and damage from leafminer and rust mites. Both simple layer and tree-type classification models based on SVM were developed.

This study attempted to provide a recognition and classification method for typical citrus surface blemishes to improve fruit quality and operation efficiency of citrus packinghouses and juice processing