Computers and Electronics in Agriculture 152 (2018) 64–73

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

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

Original papers

# Detection of orchard citrus fruits using a monocular machine vision-based method for automatic fruit picking applications

J.J. Zhuang[a,1], S.M. Luo[a,1], C.J. Hou[a], Y. Tang[a,*], Y. He[b], X.Y. Xue[c]

[a] College of Automation, Zhongkai University of Agriculture and Engineering, Guangzhou 510225, China
[b] College of Biosystems Engineering and Food Science, Zhejiang University, Hangzhou 310058, China
[c] Nanjing Research Institute for Agricultural Mechanization, Ministry of Agriculture, Nanjing 210014, China

ARTICLE INFO

*Keywords:*
Citrus detection
Harvesting
Automatic fruit picking
Monocular machine vision

ABSTRACT

Due to the variable illumination conditions and occlusion produced by neighbouring fruits and other background participants, vision systems are important in accurately and reliably detecting mature citrus in natural orchard environments for automatic fruit picking applications. A robust citrus fruit detection method based on a monocular vision system was proposed. An adaptive enhanced red and green chromatic map was generated from an illumination-compensated image, which was obtained using block-based local homomorphic filtering. Otsu thresholding, morphology operation, marker-controlled watershed transform and convex hull operation methods were then used in combination to locate potential citrus regions from the chromatic map. Local texture information was extracted from the potential regions using local binary patterns and fed to a histogram intersection kernel-based support vector machine to make the final decision. The performance of the proposed method was evaluated on 127 test images captured in two citrus orchards on both sunny and cloudy days. Under strict PASCAL criteria, the recall rate of correctly detected citrus was greater than 0.86, with 13 false detections.

## 1. Introduction

With the increasing costs of orchard management and decreasing availability of skilled labour, harvesting citrus using traditional labour-intensive methods has become less sustainable (Gongal et al., 2015). Robotic fruit picking equipment, composed of a mobile platform, manipulator, end-effector, fruit detection system and control system, is an automated device used for harvesting (Wang et al., 2017). The fruit detection system, designed to recognize and locate fruits accurately, plays a vital role in automatic fruit picking applications. Numerous studies (Wang et al., 2017; Xu and Lv, 2018; Dorj et al., 2017; Bulanon et al., 2008, 2009; Song et al., 2014; Wachs et al., 2010) have discussed the implementation of a fruit detection system based on machine vision. The major procedures (Gongal et al., 2015) involved in fruit detection systems are image acquisition, image pre-processing, potential fruit segmentation, image post-processing, feature extraction, and/or fruit classification.

Different sensing technologies can influence the complexity and implementation difficulty of the algorithms involved in the vision-based detection systems. Hyperspectral imaging (Okamoto and Lee, 2009) provides abundant information about fruit regions, but it is time consuming to extract informative features from redundant information in much larger data spaces. Although stereo and multi-imaging techniques (Mehta et al., 2017; Mehta and Burks, 2016) can provide multi-view and depth information for objects, high-level precision is required to conduct registration to fuse multi-view images, which increases the complexity of the configuration of the vision system. In addition, most growers and farmers in China prefer lower cost and lower complexity assistant devices for agricultural operations (Zhao et al., 2016). Previous studies (Xu and Lv, 2018; Zhao et al., 2016; Lu and Sang, 2015; Sengupta and Lee, 2014) show that monocular vision systems also obtain acceptable accuracy in detecting fruits within the tree canopy. Therefore, this study focuses on a vision system using a monocular camera in the visible spectrum due to its accessibility and low cost.

Xu and Lv (2018) adopted a global monomorphic filter to conduct image enhancement and obtained good results on images with uneven illumination distribution. Wang et al. (2017) proposed an improved wavelet transform and Retinex-based method to highlight images captured from various illumination scenarios in which the potential fruit regions were segmented using *k*-means clustering. A similar clustering method for locating regions of interest was also adopted by Wang et al. (2016). Zhao et al. (2016) proposed a fusion method based on an

* Corresponding author.
*E-mail address:* tangyu_mycauc@163.com (Y. Tang).
[1] The first two authors contributed equally to this manuscript and should be considered co-first authors.

https://doi.org/10.1016/j.compag.2018.07.004
Received 2 February 2018; Received in revised form 29 June 2018; Accepted 2 July 2018

**Abbreviations**

| | |
|---|---|
| RGB | Red/Green/Blue colour space |
| AERGCM | Adaptive Enhanced Red and Green Chromatic Mapping |
| MCWT | Marker-Controlled Watershed Transform |
| LBP | Local Binary Patterns |
| SVM | Support Vector Machine |
| HSV | Hue/Saturation/Value colour space |
| BLHF | Block-based Local Homomorphic Filtering |
| SROIs | Separate Regions of Interest |
| HIK | Histogram Intersection Kernel |
| ROC | Receiver Operating Characteristics |

adaptive red/blue (RB) chromatic map and the sum of the absolute transformed difference to segment potential citrus regions and adopted a support vector machine (SVM) to recognize citrus based on several optimal co-occurrence matrix features. Lu and Sang (2015) combined RB chromatic aberration information and normalized a red/green/blue (RGB) model to segment mature citrus; they used the Canny operator to extract contour fragments of the segmented regions and recognized occluded citrus by ellipse fitting from the contour fragments. Based on the sliding window scanning technique, Kurtulmus et al. (2011) proposed a method fusing colour information, circular Gabor texture analysis and 'eigenfruit' to detect citrus. Sengupta and Lee (2014) combined shape analysis, Canny edge detection, the Hough transform, texture-based SVM classification and the scale-invariant feature transform (SIFT) into a majority voting scheme to detect and locate citrus fruits. Recently, Chen et al. (2017) proposed a data-driven fruit detection and counting method based on a deep learning strategy that could be generalized across variable unstructured environments. Although this deep learning strategy achieves better performance gains for fruit detection, its implementation depends heavily on the adopted hardware devices; for example, Chen et al. (2017) trained their network based on the NVIDIA Titan X GPU, and more than 50 thousand iterations were required for the model to converge.

Despite much effort over the last ten years, the detection of citrus fruits using accessible vision systems accurately and reliably remains challenging for two main reasons: (1) the uncertain and variable illumination conditions in orchard environments, which can decrease the quality of in-field imaging (Wang et al., 2017), and (2) the complicated canopy structures of growing fruits (Gongal et al., 2015), e.g., the occlusion of fruits in resultant canopy images by branches, leaves and other fruits can substantially limit the detection accuracy.

The overall objective of this study is to detect mature citrus fruits in orchard environments with variable illumination conditions and occlusion. The specific objectives are three-fold: (1) to propose block-based local homomorphic filtering (BLHF) and adaptive enhanced red/green chromatic mapping (AERGCM) methods to automatically improve the contrast between the citrus and background regions for images with various illumination distributions, (2) to develop a combination of the marker-control watershed transform (MCWT) and convex hull methods to extract potential citrus occluded by background participants or other citrus, and (3) to filter out false detections using local binary patterns (LBP) and a histogram intersection kernel-based SVM.

## 2. Materials and methods

### 2.1. Experimental data

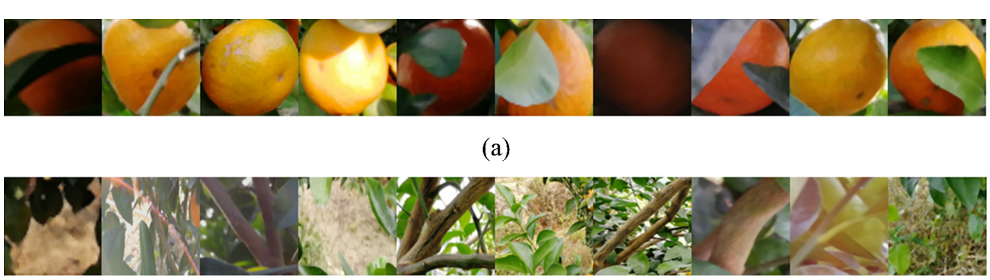
Mandarins (*Citrus reticulata* Blanco 'Shantanju') were investigated in collaboration with Guangzhou Conghua Hualong Fruit & Vegetable Freshness Co. Ltd. To evaluate the proposed citrus detection method, 227 images were captured using a camera mounted on a Huawei Honor 8 mobile phone in two natural citrus orchards approximately 3 weeks before harvest in late December 2017. With a resized resolution of 480 × 640 pixels, all images were collected with the camera located approximately 40–100 cm from the citrus fruits on both sunny and cloudy days. In total, 100 images were randomly selected to generate the training dataset, and the remaining 127 images formed the test dataset. A total of 778 samples (including 458 citrus samples and 320 non-citrus samples) with an average size of 95 × 103 pixels were extracted manually from the training dataset and used to train the citrus classifier. All the training samples were resized to 100 × 100 pixels before feature extraction, with some examples shown in Fig. 1. The citrus samples were collected under variable illumination and occlusion conditions, and the non-citrus samples were randomly collected from those background participants representing high-level differences in red and green intensities. Citrus fruits in the test dataset were manually annotated with tight bounding boxes and used for generating ground truth.

### 2.2. The proposed detection method

Fig. 2 summarizes the four main procedures of the proposed citrus detection method: local image illumination enhancement, foreground region segmentation, overlapped and occluded foreground region extraction and citrus fruit recognition.

#### 2.2.1. Local image illumination enhancement

The surface characteristics of citrus can be affected by non-uniform illumination conditions during the data acquisition phase. In addition, "over-exposure" might be caused by traditional homomorphic filtering when processing images with a uniform illumination distribution. Thus, a local image illumination enhancement method termed BLHF, was proposed. The RGB colour space of the input image, $I_{RGB}$, is first converted to a hue/saturation/value (HSV) colour space. To prevent the hue and saturation of the citrus fruits from being altered, only the intensity component $V$ is extracted for illumination enhancement.



(a)

(b)

**Fig. 1.** Examples of training samples at the resized resolution of 100 × 100 pixels: (a) citrus samples and (b) non-citrus samples.