Computers and Mathematics with Applications 64 (2012) 739–746

Contents lists available at SciVerse ScienceDirect

Computers and Mathematics with Applications

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

Detection and classification of areca nuts with machine vision

Kuo-Yi Huang *

Department of Bio-Industrial Mechatronics Engineering, National Chung Hsing University, Tai-Chung 402, Taiwan

ARTICLE INFO

Article history:
Received 18 October 2011
Received in revised form 21 November 2011
Accepted 22 November 2011

Keywords:
Areca nut
Neural network
Machine vision

ABSTRACT

In this study, we present an application of neural networks and image processing techniques for detecting and classifying the quality of areca nuts. Defects with diseases or insects of areca nuts were segmented by a detection line (DL) method. Six geometric features (i.e., the principle axis length, the secondary axis length, axis number, area, perimeter and compactness of the areca nut image), 3 color features (i.e., the mean gray level of an areca nut image on the R, G, and B bands), and defects area were used in the classification procedure. A back-propagation neural network classifier was employed to sort the quality of areca nuts. The methodology presented herein effectively works for classifying areca nuts to an accuracy of 90.9%.

1. Introduction

Areca nut is a popular and important crop for some Taiwanese. The output value of areca nut outstrips NTD 100,000 million each year in Taiwan. Though chewing areca nut is the cause of oral cancer [1], throughout Taiwan, thus the areca nut is usually called Taiwanese gum. Areca is typically cultivated on mountainsides with mild air circulation. Areca nut is often infected by various pathogens including fungi, bacteria, viruses, or harmful insects. If the surface of areca nut is damaged, the price will go down. So far it has only been sorted by traditional manpower in Taiwan. The manual cost and sorting time always impacts the income of the farmers.

Image processing is a powerful tool and widely applied to detect the agricultural products. Color, geometric and texture features are often used to analyze the quality of images. Park et al. [2] proposed a method of content-based image classification using a neural network with texture features, such as contrast, diagonal moment, energy, entropy, homogeneity, second diagonal moment, and uniformity, as input nodes. Hsieh et al. [3] used a neural network to recognize the growth stage of head cabbage seedlings according to nine texture features. Guyer and Yang [4] employed genetic artificial neural networks and spectral imaging to detect defects for cherries. Indeed, neural networks, color and texture features were often employed in the classification of plants and crops [5–12]. Huang and Lin [13] applied the boundary-chain-code, Hotelling transformation, the golden section search method, and Bayesian classification to develop algorithms to estimate the geometric characteristics of Phalaenopsis seedlings. Those geometric characteristics were executed in the sorting process. Camargo and Smith [14] used color transformation, image enhancement and a located optimum threshold to segment banana leaf infected with diseases. Shen et al. [15] employed the Otsu method, the HSI color system, and Sobel operator to extract disease spot regions and calculate leaf areas. Park et al. [16] proposed an unsupervised segmentation algorithm with Gaussian mixture models to segment color image regions. Chen and Chou [17] used discrete wavelet transform to detect Mura defects in an LCM panel of TFT-LCD.

Currently, the labor cost has risen to up to 70% of the total cost for the inspection of blemished areca nuts in Taiwan. Thus, farmers face the problem of excessive costs, and should improve their sorting process as soon as possible by a newly

* Tel.: +886 4 22840377x508.
E-mail address: kuoyi@dragon.nchu.edu.tw.

0898-1221/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.camwa.2011.11.041

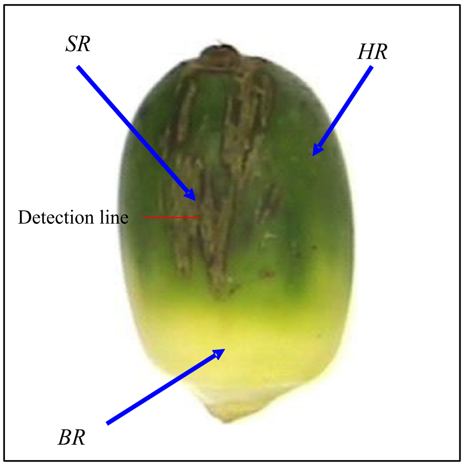

Fig. 1. An areca nut image.

developed automatic detecting and classifying system. Therefore, the aim of this study is to design a machine vision system to detect and classify the areca nut blemished with disease or insects. The technical goals are to develop an algorithm to extract the color features, geometric features, and the defect area of areca nuts; and then to classify different grades by using the aforementioned features.

2. Materials and methods

2.1. Image acquisition system

A machine vision system was developed to capture areca nut images. This system includes a GigE CCD (coupled-charge device) color camera (DFK-31AG03, Imaging Source Inc.) with a zoom lens, and a personal computer (Intel Pentium 4 processor 2.4 GHz). Open Source Computer Vision Library (OpenCV 1.0, Intel Corporation) was linked to the programs to get RGB color images of 640 $\times$ 480 pixels. The CCD camera was employed for image acquisition with 4600 lx and F4.0 opening (iris diaphragm). Images were stored in the hard drive of a PC in tagged image file (TIF) format. Image processing was executed using Microsoft Visual C++ 6.0.

2.2. Features extraction

Once the features of the areca nut have been extracted, segmentation of the entire image of areca nut is an essential procedure. The entire image of areca nut is segmented by the thresholding, hole-filling, closing, and opening operations [18]. Firstly, the principle axis of the areca nut has to be identified. By assuming the binary image of areca nut is $f(x_i, y_i)$ (where $i = 1, 2, \ldots, m$, and the total number of pixels is m), the centroid is obtained as $\bar{X} = \sum_{i=1}^{m} x_i/m$, $\bar{Y} = \sum_{i=1}^{m} y_i/m$. The covariance matrix is defined as $C = (\sum_{i=1}^{m} U_i U_i^T)/m - MM^T$, in which U_i is the ith coordinate vector of the image and $M = (\sum_{i=1}^{m} U_i)/m$ is the mean vector. T indicates vector transposition. A pair of orthogonal eigenvectors of the covariance matrix is calculated. The geometric features—the principle axis length (L_p), secondary axis (L_s), the centroid, axis number (L_p/L_s), area (A), perimeter (P), compactness ($4\pi A/P^2$), and color features—R_m, G_m, and B_m (i.e., the mean gray level of areca nut on the R, G, and B bands) of the entire areca nut are computed using eigenvectors. Secondly, these geometric and color features will be employed in the areca nut classification process.

2.3. Spot region detection

There are healthy regions (*HR*), base regions (*BR*), and spot regions (*SR*) in the appearance of areca nut, as shown in Fig. 1. Segmenting *SR* effectively is important once the harmed areca nut has been detected. However, the difference between *SR* and *HR* is difficult to be distinguished with the thresholding method. For example, the decision rule of thresholding cannot determinate to extract *SR* according to the gray level histogram. In Fig. 2, the threshold values (T) 160, 150 and 80, are used to segment *SR* on red (R), green (G) and blue (B) bands (Fig. 2(a)–(c)) individually. The segmentation results were shown in Fig. 2(g)–(i). Therefore, a novel method had to be proposed for *SR* segmentation of areca nut in the study.

A prior experiment was preceded as follows. Firstly, a detection line with scanning resolution in a pixel was employed to find gray levels of areca nut image. There are different distribution forms of gray level in *HR*, *SR* and *BR* as illustrated in Fig. 3. The curves of gray level on R and G bands are quite close on the region of *HR*, as shown in Fig. 3(a). Even some