Expert Systems with Applications 39 (2012) 780–785

Contents lists available at ScienceDirect

Expert Systems with Applications

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

Detecting rottenness caused by *Penicillium* genus fungi in citrus fruits using machine learning techniques ☆

Juan Gómez-Sanchis [a,*], José D. Martín-Guerrero [a], Emilio Soria-Olivas [a], Marcelino Martínez-Sober [a], Rafael Magdalena-Benedito [a], José Blasco [b]

[a] *Electronic Engineering Department, E.T.S.E, University of Valencia, C/Dr Moliner 50, 46100 Burjassot, Valencia, Spain*
[b] *Centro de Agroingeniería. IVIA, Ctra. Moncada-Náquera, km. 5, 46113 Moncada, Valencia, Spain*

ARTICLE INFO

Keywords:
Hyperspectral imagery
Rottenness
Citrus fruits
Neural networks
CART
Agriculture

ABSTRACT

Penicillium fungi are among the main defects that may affect the commercialization of citrus fruits. Economic losses in fruit production may become enormous if an early detection of that kind of fungi is not carried out. That early detection is usually based either on UltraViolet light carried out manually. This work presents a new approach based on hyperspectral imagery for defect segmentation. Both the physical device and the data processing (geometric corrections and band selection) are presented. Achieved results using classifiers based on Artificial Neural Networks and Decision Trees show an accuracy around 98%; it shows up the suitability of the proposed approach.

1. Introduction

Citrus industry involves a huge amount of money worldwide. There are many factors that may modify citrus production, being the presence of fungi one of the most relevant ones. In particular, the losses due to the fungi *Penicillium italicum* and *Penicillium digitatum* are especially dramatic (Eckert & Eaks, 1989). The detection of those fungi in a fruit production is crucial in order to use specific treatments that can avoid their propagation over all the production. Nowadays, rottenness detection is carried out manually by one method. This method is based on using UltraViolet (UV) light which produced a fluorescence in the affected area; nonetheless, there are some problems associated with this method, namely, the UV light is harmful to the worker and the procedure cannot distinguish which kind of fungus is responsible for the damage. Generally, RGB computer vision systems are used for the automation of classification processes based on color or shape as well as detection defects that are observable at first glance (Blasco, Aleixos, & Moltó, 2003, 2007; Blasco et al., 2007; Leemans, Magein, & Destain, 2002); however, they cannot classify other defects that are not directly observable, such as rottenness produced by *Penicillium* fungi.

Advanced multispectral computer vision systems are proposed as a solution to find this kind of defects while preserving workers safety (Aleixos, Blasco, Navarrón, & Moltó, 2002; Kleynen, Leemans, & Destain, 2005). Multispectral and hyperspectral systems are characterized by the use of spectral information beyond the three classical bands: right, green and blue. Moreover, they are not limited to the visible spectrum but also to other areas, like the Near InfraRed (NIR). However, the number of used bands must be reduced since a great number of bands would involve a performance decrease due to an excessive time for acquiring a whole hyperspectral image. An additional problem related to the high dimensionality of this kind of imagery is the worsening of classifiers performance when the number of input features is high (Friedman, 1994; Hastie, Tibshirani, & Friedman, 2001). Therefore, a key step in the proposed expert system is the choice of those hyperspectral bands that have the highest discriminating ability in the tackled problem. Once the most relevant bands are selected, the next step is the use of different classifiers; in this work, we focus on Classification And Regression Trees (CART) and Artificial Neural Networks (ANN), more specifically the MultiLayer Perceptron (MLP).

The rest of paper is outlined as follows. Section 2 shows the computer vision system used for the acquisition of hyperspectral imagery. In Section 3, labeled data sets are produced with the two following goals: (1) description of the spectral characteristics of the classes defined by an expert from the images acquired by the system presented in Section 2; and (2) development and evaluation of the models proposed for rottenness detection. Section 4 presents the band selection and the description of the classifiers (CART and MLP). Achieved results are shown in Section 5, ending up the paper with the conclusions of the work in Section 6.

☆ This work was partially funded by the Spanish Ministerio de Ciencia e Innovación (MICINN) through research Project DPI2010-19457, with the support of European FEDER funds.

* Corresponding author.
E-mail address: juan.gomez-sanchis@uv.es (J. Gómez-Sanchis).

0957-4174/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2011.07.073

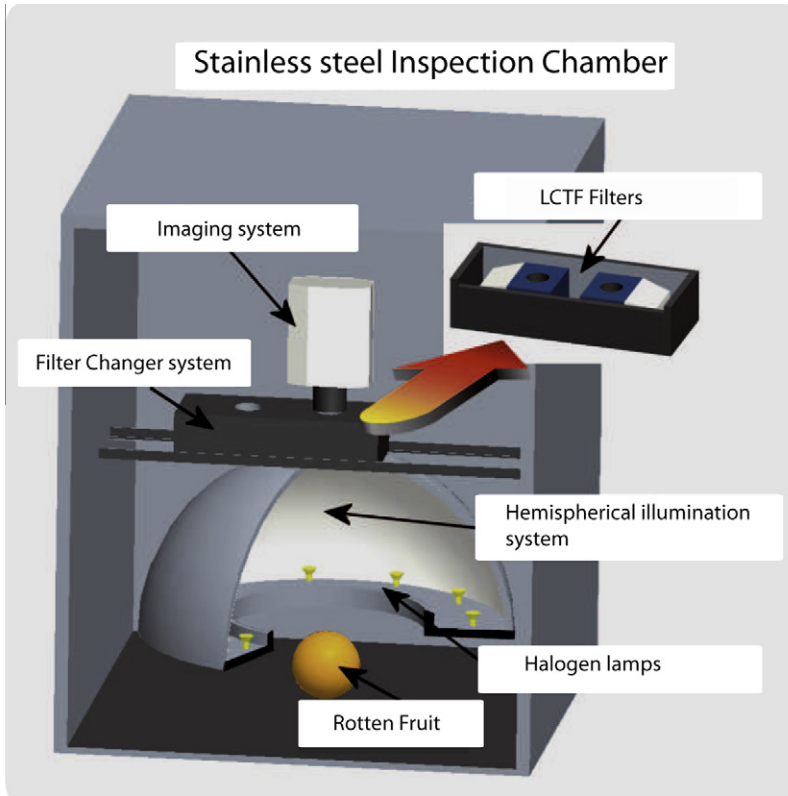

Fig. 1. Hyperspectral vision system.

2. Imaging system

Image acquisition and frequence selection of the hyperspectral vision system was based on a high-resolution monochromatic camera[1] with 4.32 pixels/mm, Schneider's Xenoplan optics (1.4/17 mm) with low chromatic dispersion and two liquid crystal tunable filters (LCTF) (Varispec VIS-07 and NIR-07); both filters have a spectral resolution of 10 nm, the first filter ranges from 400 nm to 720 nm while the second is sensitive from 650 nm up to 1100 nm. The visible part of the image was acquired by the VIS-07 filter whereas the NIR-07 filter dealt with the NIR spectrum. Fig. 1 shows the system for filter exchange developed by the authors. It can acquire hyperspectral imagery systematically.

The acquired hyperspectral images were transferred to a computer[2] by means of an acquisition card property of the camera. The acquisition system was focused on the central band of the acquisition range (740 nm).

Each fruit was lighted up by an indirect system formed of 12 halogen lamps of 20 W arranged in an aluminium hemispheric diffuser. The lighting system was powered by a power supply unit (12 V and 350 W). The so-called integration time (time needed for acquiring a monochromatic image from the hyperspectral one) was calculated to make the system work in the central area of its dynamic range with a flat spectral response (on average). It was ensured using a 99% certified white reference. Fig. 1 shows how all the hyperspectral vision system was inside a stainless steel closed unit to avoid negative interferences from the light present in the acquisition room.

[1] Coolsnap ES model, Photometrics.

[2] The processor used in this work was an Intel Core Duo 1.67 GHz, with RAM memory of 2 GB.

3. Data sets and data mining process

3.1. Vegetal material

The vegetal material that was used to build the labeled data sets was formed by tangerines (Citrus clementina Hort. ex Tanaka) with two kinds of damages:

- Common damages that are time-invariant (produced by *trips* and branch frictions). They are visible at first glance. Those fruits affected by this kind of damages were selected randomly in an agricultural products company.
- Damages that do change with time (produced by the fungi *P. digitatum* and *P. italicum*). At the beginning, these damages presented a color similar to normal fruit, hence they could not be detected in a routine visual inspection. In this work, this kind of damages was produced by inoculating the fruit with spores of the fungi.

The data set was formed by 120 fruits in different states of ripening, 30 of which did not present any defect, 30 presented defects by *trips* or branch frictions, (both defects are very similar and produce a similar depreciation of the fruit), 30 were inoculated with spores of *P. digitatum* and the remaining 30 fruits were inoculated with spores of *P. italicum*. The inoculation was based on a dissolution of spores in suspension with a concentration of 10^6 spores/ml. This concentration is usually used to produce rottenness (Palou, Smilanick, Usall, & Viòas, 2001). Fruits were preserved in a controlled atmosphere of 25 °C and relative humidity of 99% for three days. After that period, all tangerines inoculated by the fungi presented damages in an early stage of formation with a variable diameter ranging from 15 mm to 40 mm. Hyperspectral images were acquired for every fruit (120 images), in a wavelength