BIOSYSTEMS ENGINEERING 122 (2014) 91–98

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/issn/15375110

Research Paper

Early detection of mechanical damage in mango using NIR hyperspectral images and machine learning

Nayeli Vélez Rivera [a], Juan Gómez-Sanchis [b], Jorge Chanona-Pérez [a], Juan José Carrasco [b], Mónica Millán-Giraldo [b,c], Delia Lorente [d], Sergio Cubero [d], José Blasco [d,*]

[a] Instituto Politécnico Nacional — Escuela Nacional de Ciencias Biológicas, Departamento de Ingeniería Bioquímica, Av. Plan de Ayala y Carpio s/n, Col. Santo Tomás, CP 11340, Mexico, D.F., Mexico

[b] Universitat de València — Intelligent Data Analysis Laboratory, Av. Universitat S/N, 46100, Burjassot, Valencia, Spain

[c] Universitat Jaume I — Institute of New Imaging Technologies, Av. Sos Baynat S/N, 12071, Castelló de la Plana, Spain

[d] Centro de Agroingeniería, Instituto Valenciano de Investigaciones Agrarias (IVIA), Ctra. Moncada-Náquera km 5, 46113, Moncada, Valencia, Spain

ARTICLE INFO

Article history:
Received 26 September 2013
Received in revised form
19 February 2014
Accepted 27 March 2014
Published online 8 May 2014

Keywords:
Computer vision
Feature selection
Fruit quality
Mango fruits
Non-destructive inspection
Hyperspectral imaging

Mango fruit are sensitive and can easily develop brown spots after suffering mechanical stress during postharvest handling, transport and marketing. The manual inspection of this fruit used today cannot detect the damage in very early stages of maturity and to date no automatic tool capable of such detection has been developed, since current systems based on machine vision only detect very visible damage. The application of hyperspectral imaging to the postharvest quality inspection of fruit is relatively recent and research is still underway to find a method of estimating internal properties or detecting invisible damage. This work describes a new system to evaluate mechanically induced damage in the pericarp of 'Manila' mangos at different stages of ripeness based on the analysis of hyperspectral images. Images of damaged and intact areas of mangos were acquired in the range 650–1100 nm using a hyperspectral computer vision system and then analysed to select the most discriminating wavelengths for distinguishing and classifying the two zones. Eleven feature-selection methods were used and compared to determine the wavelengths, while another five classification methods were used to segment the resulting multispectral images and classify the skin of the mangos as sound or damaged. A 97.9% rate of correct classification of pixels was achieved on the third day after the damage had been caused using *k*-Nearest Neighbours and the whole spectra and the figure dropped to 91.4% when only the most discriminant bands were used.

* *Corresponding author*. Tel.: +34 96 3424000.
E-mail addresses: jchanona@ipn.mx (J. Chanona-Pérez), blasco_josiva@gva.es (J. Blasco).
http://dx.doi.org/10.1016/j.biosystemseng.2014.03.009

1. Introduction

The mango (*Mangifera indica* L.) has become important as a product to be exported for several producer countries. The FAO reported that in 2011 Mexico was ranked 5th among top mango-producing countries and was also the world's number one exporter (FAOSTAT, 2011). Mexico produces more than 50 varieties of mango, 'Manila' being one of the most important for both the national and international markets (Ornelas-Paz, Yahiaa, & Gardea, 2008). The 'Manila' mango is a very soft fruit and it is therefore very susceptible to mechanical damage during postharvest processes. Mechanical damage to fruit results from surface loads, but the factors determining the different susceptibility or sensitivity of the fruit to mechanical damage from one species or variety to another is still unknown (Martínez-Romero et al., 2004), and its quantitative evaluation is still a challenge (Li & Thomas, 2014). Hahn (1999) reported on the effect of damage caused by impact on different varieties of mango and 'Manila' was one of the most sensitive, since this cultivar began to develop bruising after being impacted from heights of only 20 cm. One of the main problems in the postharvest period is that the damage caused during postharvest handling may not be visible until several days later, when the fruit has already been marketed. Any process, and consequently the reputation of the producer, is judged by the quality of the final product, which makes quality control a crucial stage of the process. Hence, this is a big problem in the mango industry because this lack of quality cannot be detected by visual inspection during postharvest handling, and damaged fruit with a bad appearance can have an important influence on the consumer's decision to choose a different kind of fruit. However, this situation becomes even worse when the fruit is exported because this is a priority for Mexican fruit and only fruits of excellent quality can be exported.

The detection of damaged fruit in the industry can be achieved by means of automatic systems (Valiente-González, Andreu-García, Potter, & Rodas-Jordá, 2014). Non-destructive methods such as spectrometry or image analysis in the visible and near-infrared range of the spectrum have been widely applied to evaluate aspects of quality and maturity in fruits and vegetables (Cubero, Aleixos, Moltó, Gómez-Sanchis, & Blasco, 2011), including mango fruit (Vélez-Rivera et al., 2014; Wanitchang, Terdwongworakul, Wanitchang, & Nakawajana, 2011). For instance, Hahn (2004) related mechanical damage with loss of firmness in 'Kent' mangos. The author induced the impacts with a pendulum device so as to be able to control the forces involved in causing the damage. In fact, controlled impacts are frequently used in studies for evaluating mechanical damage in olives (Jiménez-Jiménez, Castro-García, Blanco-Roldán, Agüera-Vega, & Gil-Ribes, 2012), oranges (Ortiz, Blasco, Balasch, & Torregrosa, 2011), apples (ElMasry, Wang, Vigneault, Qiao, & ElSayed, 2008), persimmon (Novillo, Salvador, Llorca, Hernando, & Besada, 2014), potatoes or tomatoes (Van Zeebroeck et al., 2003).

Nevertheless, various defects like spoilage, microbial contamination, mechanical and freeze damage or fungal infections are very difficult to detect at early stages using standard artificial vision systems (Gómez-Sanchis et al., 2013) and other technologies need to be developed. The advances made in hyperspectral imaging systems, together with the reduction in their price, have recently allowed this technology to be incorporated into postharvest quality control laboratories, which are using it to develop new advanced systems to predict the internal quality and properties of fruit and vegetables (Lorente, Aleixos, et al., 2012), including some types of mechanical damage (Nanyam, Choudhary, Gupta, & Paliwal, 2012). This technology combines the analysis of both spatial and spectral characteristics of a sample (ElMasry, Kamruzzaman, Sun, & Allen, 2012). As a result, it can therefore make use of the spectral information to estimate the internal compounds or also to detect diverse types of damage that cannot be detected by standard computer vision systems because the damage is not fully developed or its appearance is very similar to the sound skin, as is the case of decay lesions in citrus fruits (Blasco, Aleixos, Gómez-Sanchis, & Moltó, 2009) or internal properties in bell peppers (Schmilovitch et al., 2014). However, the complexity of the analysis of hyperspectral images requires a great deal of computing time. In addition, the huge amount of data provided by these systems makes it necessary to have reliable methods to select only the information that can actually optimise the results. Since a hyperspectral image consists of a set of consecutive narrow band monochromatic images, some information contained in close bands is redundant and frequently highly correlated. For this reason, it is very important to select only those bands with the most relevant information, while discarding those that do not contribute in any significant way to improve the results. Under this premise several studies have recently been carried out using different feature selection methods. For instance, Lorente, Aleixos, Gómez-Sanchis, Cubero, and Blasco (2013) and Lorente, Blasco, et al. (2012) compared several methods for selecting important wavelengths to detect decay lesions in citrus fruits at early stages. Serranti, Cesare, and Bonifazi (2013) used interval partial least squares discriminant analysis (iPLS-DA) on hyperspectral images to reduce their spectral dimension and redundancy, and to extract useful image features for differentiating three classes of wheat.

Once the relevant bands have been selected, a multispectral image is obtained, which then has to be segmented in order to achieve the real performance of the system. This segmentation consists of classifying each pixel in the multispectral image as belonging to a region of interest, using the spectral value of the pixels in the different selected bands as inputs of the classifiers (Gómez-Sanchis et al., 2012, 2013). Hence, the main objective of the experiments carried out is to study the possibility of detecting mechanical damage induced intentionally in 'Manila' mangos at early stages of maturity using a hyperspectral vision system, which is achieved by identifying the spectral bands that best categorise whether a mango is damaged or not. The work also aims to obtain the classifier that best fits this problem by segmenting the images using only the selected bands.

2. Material and methods

2.1. *Fruit used in the experiments*

Ten mangos (*Mangifera indica* L.) cv. 'Manila' were collected from a local market, all of them being similar in terms of stage of maturity (unripe), firmness, colour (dark green), shape and