Journal of Food Engineering 115 (2013) 407–414

Contents lists available at SciVerse ScienceDirect

Journal of Food Engineering

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

Assessment of chestnut (*Castanea* spp.) slice quality using color images

Irwin R. Donis-González [a,*], Daniel E. Guyer [a], Gabriel A. Leiva-Valenzuela [b], James Burns [a]

[a] Department of Biosystems and Agricultural Engineering, 216 Farrall Hall, 524 S. Shaw ln., Michigan State University, East Lansing, MI 48824, United States
[b] Department of Chemical and Bioprocess Engineering, Pontifica Universidad Católica de Chile, Avenida Vincuña Mackenna 4860, Macul, Santiago, Chile

ARTICLE INFO

Article history:
Received 21 May 2012
Received in revised form 28 August 2012
Accepted 27 September 2012
Available online 23 October 2012

Keywords:
Classification
Computer vision
Pattern recognition
Processed chestnuts

ABSTRACT

Unbiased internal quality classification of Chestnuts (*Castanea* spp.) is extremely important to the fresh and processed industries. It can also be used as a tool for applied scientific studies, such as the training of non-invasive techniques to determine chestnut internal quality, and the effect of pre- and post-harvest treatments. At the moment, humans visually perform the invasive quality assessment of chestnuts. This procedure is prone to errors and high variability due to individuals' fatigue, lack of training, and subjectivity. Thus, there is a need to develop a technique that is able to objectively classify internal quality of chestnuts. In this paper, a computer vision methodology is proposed to sort chestnuts into five classes, as established by an expert human rater. 1790 color images from slices with different quality classes were acquired, using a flat panel scanner, from the hybrid cultivar 'Colossal' and 'Chinese seedlings'. After pre-processing, a total of 1931 color, textural, and geometric features were extracted from each color image. Furthermore, the most relevant features were selected using a sequential forward selection algorithm. Thirty-six features were found to be effective in designing a quadratic discriminant classifier with a cross-validated overall performance accuracy of 89.6%. These results showed that this method is an accurate, reliable, and objective tool to determine chestnut slice quality, and might be applicable to in-line sorting systems.

1. Introduction

Chestnut (*Castanea* spp.) is an agricultural commodity with increasing international interest and surges in customer consumption in Europe, Australia, New Zealand and the United States (Gold et al., 2006). For example, in 2009, the US imported approximately 4,876 T of chestnuts (http://faostat.fao.org/). The high importing of chestnuts corresponds with an US chestnut industry that is in its infancy, and has potential for growth (Gold et al., 2006; Donis-González, 2008). In the US, chestnuts are sold fresh and processed (e.g. peeled frozen, and dry slices) (http://www.chestnutgrowers-inc.com/; Donis-González, 2008). It is important to provide high quality chestnuts, and an overall satisfactory experience of both fresh and processed chestnuts, in order for the consumers to become repeat customers (Gold et al., 2006).

Current invasive standards for fresh chestnut quality are laid out in the agricultural quality standards of the United Nations Economic Commission for Europe FFV-39 (UNECE, 2010). These standards are ambiguous, highly subjective, and divide chestnuts into three classes; 'Extra', 'I', and 'II', with decreasing quality respectively. Minimum requirements are that chestnuts need to be, intact, sound, clean, free from pest damage, not germinated, and free from foreign tastes or smells. There are allowances for slight defects in color, development, and shape for the lesser class. As with many agricultural products, chestnut quality is invasively assed by humans (Brosnan and Sun, 2002; Donis-Gonzalez et al., 2012) resulting in a time-consuming and labor intensive method, which is inherently subjective to human error due to inexperience, fatigue, and distractions (Brosnan and Sun, 2002).

Computer vision offers an alternative to visual inspection, and has been used in various foods and agricultural commodity sorting systems today, being objective, consistent, rapid, and economical (Brosnan and Sun, 2002; Kumar-Patel et al., 2012). Color computer vision has been effectively used to classify or recognize quality in several agricultural and food commodities including apples (*Malus domestica*) (Paulus and Schrevens, 1999), strawberries (*Fragaria* spp.) (Bato et al., 2000), pistachios (*Pistacia vera*) (Pearson and Toyofuku, 2000), external damage induced by worms in chestnuts (Wang et al., 2011), potato chips (Pedreschi et al., 2006), tortillas (Mery et al., 2010), pizza (Sun and Brosnan, 2003a, 2003b),

Abbreviations: CCD, charge-coupled-device; CGI, chestnut Growers Inc.; CT, computer tomography; DPI, dots-per-inch; FOSMOD, forward orthogonal search algorithm maximizing the overall dependency; KNN, K-neareast neighbor; LBP(d,s), local binary patterns; LDA, linear discriminant analysis; MD, Mahalanobis distance; PNN, probabilistic neural network; QDA, quadratic discriminant analysis; RANKFS, rank key features by class sorting criteria; ROC, relative operating characteristic curve; SFS, sequential forward selection; SVM, support vector machine; TIFF, tagged image file format; *Tx*(k,p), Haralick texture; UNECE, United Nations Economic Commission.

* Corresponding author. Tel.: +1 517 432 7438; fax: +1 517 432 2892.
E-mail address: donisgon@msu.edu (I.R. Donis-González).

0260-8774/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.jfoodeng.2012.09.017

chocolate chip biscuits (Davidson et al., 2001), cheese (Wang and Sun, 2001, 2002a, 2002b), and domestic pork meat (Lu et al., 2000; Faucitano et al., 2005). However, computer vision has never been used to sort and assess internal chestnut quality, and there is no methodology that supports the invasive quality assessment of fresh and processed chestnuts, including fresh chestnut slices.

A fundamental of computer vision systems are pattern recognition algorithms. The computer vision system is trained from specific patterns of interest extracted from a set of color images (e.g. different chestnut quality categories). A pattern or feature is represented by a group of geometric and image intensity features, which are able to define all of the quality classes. The computer vision system then assigns a new image to a specific quality category or class (Duda et al., 2000). The first step consists in extracting a high number of features (patterns) from the category of known images. After that, features must be selected by their capacity of correctly separating the classes, therefor training the system, and allowing it to automatically classify a new image. Classification is done using statistical and clustering algorithms by assigning each image to its corresponding class (Duda et al., 2000; Mery and Soto, 2008). Complete information regarding statistical pattern recognition methods have been described in several publications, including Jain et al. (2000), Duda et al. (2000), Bishop (2007), and Holmström and Koistinen (2010).

This paper describes a statistical pattern recognition technique developed to objectively and consistently rate quality of fresh sliced chestnuts using color images. This objective measurement-rating tool could be used in research institutions as a ground-truthing to efficiently calibrate, and train, non-invasive fresh chestnut quality assessment methods, like the computer tomography (CT) system proposed in Donis-Gonzalez et al. (2012). Furthermore, it can be used as an inference technique to quantify the positive or negative effect of postharvest and preharvest treatments. In addition, this approach will enable the industry to invasively sample and forecast the overall quality of a whole fresh chestnut lot. Moreover, with slight modifications, the methodology could evolve for in-line use to directly classify the quality of peeled-sliced chestnuts in the processed chestnut industry. This would help the industry to automatically make the final decision of either discarding or keeping the product for consumption, enhancing final product quality.

2. Materials and methods

2.1. Sample collection and preparation

Physiologically mature chestnuts (*Castanea sativa* × *Castanea crenata*) cv. 'Colossal' and Chinese chestnut seedlings (*Castanea mollissima*), were obtained from Chestnut Growers Inc. (CGI; Grand Haven, Michigan, USA). These chestnuts were previously collected from seven commercial farms in Michigan from the 2009, 2010, and 2011 harvests. Each year, samples were immediately stored at 4 °C. Five days later, chestnuts within each species were mixed, randomized and submerged for 300 s in 75 L of room temperature distilled water containing 2700 μL/L hydrogen dioxide plus 200 μL/L peracetic acid (Storox®[1], BioSafe Systems, Glastonbury, CT, USA) with the objective of reducing external mold contamination.

All chestnuts were stored in mesh bags at 4 °C. After 90 d, slice image acquisition was conducted.

2.2. Chestnut slice image acquisition

Immediately after storage, each fresh chestnut was transversely sliced into 4 sections using a sharp hand knife. Slice sizes varied depending on chestnut size, but typically were between 5 and 7.5 mm thick. All chestnuts slices from five different samples, located side by side, were manually set on their flat side directly over the clean scanner glass (fixed focal point), avoiding the presence of controlled (e.g. external sample carrier) and uncontrolled foreign objects (e.g. other chestnut particles). Samples were scanned using a 48 bit color, 9600 × 4800 dots-per-inch (DPI) charge-coupled-device (CCD) scanner (Scan Maker S400, Microtek International Inc., China), using the ScanWizard 5 (Microtek International Inc., China) standard image acquisition software, yielding a tagged image file format (tiff) color image, with a resolution of 816 × 1123 pixels, as seen in Fig. 1. Scan mode was set to true color photo image.

Before every scan, the scanner was thoroughly cleaned, using compressed air in combination with wiping the scanning glass with delicate task wipes, which had been previously soaked in mild non-streak glass cleaner. To avoid variability between images, and to stabilize the intensity of the scanner lamp, the scanner was on for at least 15 min before scanning. It is important to mention that the scanner, which was used in this study, is internally calibrated every time it is turned on, so no calibration and/or calibration targets are required (http://support.microtek.com/product_dtl_2.phtml?prod_id=38).

2.3. Chestnut slice image segmentation

After image acquisition, each chestnut slice was automatically cropped using Matlab R2009a and its image processing toolbox (The Mathworks, Inc., Natick, MA, USA). Each individual slice color image had a resolution of 151 × 151 pixels. Image segmentation was implemented to recognize the region of interest in the image, which is the chestnut slice in each color image segmented from its background. A combination of simple thresholding (threshold level = −0.05 for values of pixels in normalized images between −1 and 1) and morphologic operations were used to segment each chestnut slice color image, following the optimized procedure for color image segmentation with a homogenous background, as described in Mery and Pedreschi (2005). The segmentation procedure can be found in the "Balu" free toolbox for pattern recognition (http://dmery.ing.puc.cl/index.php/balu/), developed by the Department of Computer Science at the Pontifical Catholic University of Chile (Santiago, Chile) under the Bim_segbalu-function. This toolbox contains more than 200 functions for image processing, feature extraction, feature transformation, feature analysis, feature selection, classification, clustering, performance evaluation, image sequence processing, and more.

2.4. Feature extraction and selection

Color components were extracted from color images of each slice resulting in red, green, and blue (RGB), hue saturation value (HSV), and lightness/color components (L*a*b*), using the method proposed by León et al. (2006). In addition, a gray scale image was obtained from each color image (Shapiro and Stockman, 2001). Therefore, ten intensity images were obtained from each chestnut color slice. From these ten images, 1931 features were extracted, as seen in Fig. 1. Features were extracted from each of the ten intensity images using the "Balu" toolbox. Extracted features included standard features, invariant shape moments, Haralick textural features (*Tx*), local binary patterns (LBP), and Gabor filters.

2.4.1. Standard features

Six standard features, describing simple intensity information, were derived from the segmented image region, for all of the gray scale images (Shapiro and Stockman, 2001). Standard features included the mean (μ) – Eq. (1), standard deviation (σ) – Eq. (2), kurtosis (k) – Eq. (3), skewness (s) – Eq. (4), mean gradient (first-order

[1] Storox® is a registered Trademark BioSafe Systems (Glastonbury, CT, USA).