Computers and Electronics in Agriculture 118 (2015) 340–349

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

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

A non-destructive oil palm ripeness recognition system using relative entropy

Attaphongse Taparugssanagorn [a,b,*], Siwaruk Siwamogsatham [b], Carlos Pomalaza-Ráez [c]

[a] School of Engineering and Technology/Telecommunications, Asian Institute of Technology, Thailand
[b] National Electronics and Computer Technology Center, Thailand
[c] College of Engineering, Technology, and Computer Science, Department of Engineering, Purdue University, IN, USA

ARTICLE INFO

Article history:
Received 30 July 2014
Received in revised form 4 February 2015
Accepted 18 September 2015
Available online 3 October 2015

Keywords:
Histogram color discrimination
Information theory
Kullback–Leibler divergence
Nigrescens fruits
Virescens fruits

ABSTRACT

This paper introduces a relative entropy based image processing approach for the non-destructive prediction of the maturity of oil palm fresh fruit bunches (FFB) which enables the determination of the correct time for harvesting. The results of an experimental study of applying the Kullback–Leibler distance to the problem of oil palm classification are presented. It is shown that the proposed algorithm has an excellent accuracy and it can be computed very fast. The overall proposed system is simple and useful for oil palm farmers and entrepreneurs.

1. Introduction

Oil palm (Elaeis guineensis) is cultivated on approximately 15 million hectares across the world due to its economic importance (Fitzherbert et al., 2008). It is a highly efficient oil producer, with each fruit containing about 50% of oil. As a result it requires ten times less land than other oil producing crops. Palm oil is a vegetable oil used for a variety of applications, not only edible products, such as cooking oils, margarines, baked goods, but also for non-edible purposes, such as soaps, washing powders, cosmetics, and bio-fuels. A recent trend has been the increasing use of bio-fuels to reduce the reliance on fossil fuels. This trend has created a demand for palm oil as a feedstock ingredient in the production of a biodiesel, which is also known as palm oil methyl ester, reducing diesel use and consequently reducing CO_2 emissions (Fitzherbert et al., 2008).

In the area of harvesting discipline and quality control for oil palm fruits, color has been an important guide to determine whether the oil content has reached a maximum such that the fruit bunch is ready for cutting (Ng, 1957). Two main indicators, i.e., oil extraction rate (OER) and free fatty acid level (FFA), can indirectly influence the profitability of any plantation enterprise. The national standard stipulates that the theoretical extraction rate is between 21–23% and the FFA content should not be higher than 5% (Ng, 1957). It has been recognized that this is the result of processing poor quality fresh fruit bunches (FFB). In fact for every 1% of unripe bunches present, the OER will decrease by 0.13%, while the FFA content will increase linearly as the percentage of overripe bunches increases (Siregar, 1976).

Ripeness assessment and FFB quality control in the oil palm industry is crucial for the evaluation of the processing results (Wood et al., 1984). In this respect perception of color is very important for indicating fruit maturity and defects. In nature, an oil palm tree continuously produces FFBs and the amount of oil in its fruits reaches a maximum, after which it declines due to the hydrolysis of fat and synthesis of free fatty acid (Southworth, 1976). Mature bunches detach fruit progressively and ultimately become rotten and moldy before breaking-up (Southworth, 1976). Therefore, oil palm FFBs need to be harvested at the optimum maturity. Traditionally, the ripeness level is defined in terms of the number of detached fruits from the bunch. Two different criteria for checking the number of detached fruits were introduced in Hitam and Yusof (2000). The former is the number of detached fruits on the ground before the FFB is cut, and the latter is the number of detached fruit sockets on the bunch. The former method is often used for tall trees, which has been used until today, while the latter is suitable only for short trees.

However, these methods are inaccurate, time consuming and laborious, which lead to higher harvesting and production costs.

* Corresponding author at: School of Engineering and Technology/Telecommunications, Asian Institute of Technology, P.O. Box 4, Klong Luang, Pathumthani 12120, Thailand.

E-mail address: attaphongset@ait.asia (A. Taparugssanagorn).

http://dx.doi.org/10.1016/j.compag.2015.09.018
0168-1699/

In addition, there is not really a common definition to specify which bunch is ripe or unripe (Jalil, 1994; Siregar, 1976; Southworth, 1976). Because of the problems of traditional methods, more consistent image processing techniques have been proposed for replacing these traditional methods. The correlation between the color of oil palm fruits and their oil content has been investigated showing that there is a positive correlation between both attributes, i.e., unripe fruit has the lowest oil content, ripe fruit has the highest oil content, and the oil content deteriorates when the fruit reached the overripe stage (Choong et al., 2006; Alfatni et al., 2006). Computer-based technologies and tools for machine vision, which can mimic human color recognition, have been introduced (Abdullah et al., 2012, 2004; Blasco et al., 2003; Ishak and Hudzari, 2010). However, both hardware components such as personal computers, color frame grabber, and charge-couple-device (CCD) cameras, and software for processing and control are required to provide a complete system, which might not be practical in an actual application. Patter recognition techniques using principal component analysis (PCA) has been proposed to identify different ripeness classes of oil palm FFB (Zhang and Wu, 2011). Three features represented by three RGB values are analyzed to obtain a plot of the principal components. The centroid values is then identified and used for indicating each ripeness class. Finally, the Euclidean distances between the centroid values and the plot of other samples are used to classify the oil palm FFB. This method yielded 75% correct classification for RGB images. A photogrammetric based approach, which correlates the color of the palm oil fruits to their ripeness, is presented in Ahmed et al. (2009). The methodology consists of both a hardware component in the form of an illumination chamber and a software component that calculates the color digital numbers (DN) and classifies the ripeness of the oil palm FFB. However, this method still has difficulty in differentiating the average RGB values for the ripe and unripe FFB since there is no clear-cut distinction of the RGB values for both cases. More sophisticated oil palm FFB grading system using neuro-fuzzy and fuzzy logic are developed in Jamil et al. (2009), May and Amaran (2011). Better accuracy, but more complicated techniques, like artificial neural network (ANN) classifier have been studied for various classification tasks of different agricultural products (Chinchuluun et al., 2009).

This paper proposes a simple non-destructive oil palm ripeness recognition using image processing techniques together with information theory. Images can be taken in a natural light environment. The differences of the distributions of a testing image of palm bunches and of standard scale images, which predefine each level of oil palm ripeness, are computed in terms of relative entropy, also called "Kullback–Leibler distance (KL distance)". The level of the standard scale image, which has the smallest distance from the one of the testing image, is used to decide about the ripeness level of the tested oil palm.

The rest of the paper is organized as follows. Section 2 describes the methodology of this work. All the process including the digital image pre-processing, the information-theoretical color similarity measures, the proposed technique, and the mobile application development are included in this section. Section 3 shows the experimental results. Practical issues are discussed in Section 4. The conclusions are given in Section 5.

2. Materials and methods

2.1. Palm oil grading

As previously stated, two parameters (OER and FFA) indicate the quality of fresh fruit FFB. They can be used to recognize the maturity or ripeness for grading and harvesting process of oil palm. In manual grading of oil palm, the color of FFB is the most important indicator for farmers to determine the ripeness of the oil palm fruit or FFB. Therefore, there should be a relationship between the OER with the color of FFB.

The relationship of Hue values, which is a term to describe the pure spectrum colors, of FFB images with the OER or the mesocarp oil content was developed in Razali et al. (2009, 2012). In these studies the FFB images captured are analyzed using their RGB values. Then, the RGB values are converted to the corresponding Hue digital values. In the next step the OER representing the oil content in the mesocarp parts is calculated. The whole dry weight model is used to calculate oil-to-dry mesocarp ratio (Razali et al., 2009, 2012). The Hue digital values is finally correlated to the mesocarp oil content using regression models as shown in Figs. 6 and 8 in Razali et al. (2009, 2012). The day of harvesting or a number of days before harvest of FFB was also predicted in Razali et al. (2009, 2012). The harvesting days are determined based on the 75% mesocarp oil content which indicated as a ripe stage for FFB (Razali et al., 2012). A linear interpolation technique is used to fix the date for the oil content of mesocarp reaching at 75%. At last, the harvesting days of FFB are determined by the percentage of mesocarp oil content using linear equations as shown in Fig. 14 in Razali et al.. (2012). As a result, the oil is found to start developing in mesocarp fruit at 65 days before fruit at ripe maturity stage of 75% oil-to-dry mesocarp.

In terms of palm oil quality, the FFA is an important parameter. Since fats and oils contain some level of free fatty acid, there is always an increase in acidity with time during transport and storage. The effect of light of different colors on the FFA values of stored crude palm oil was studied in Oyem (2011). In general, the effect of light on stored palm oil is that of increasing not only the rate of oxidation, but also that of hydrolysis since light is a source of energy. Therefore, it is suggested that palm oil samples for storage should be kept to inhibit the effects of light (Oyem, 2011). The change in FFB's color upon ripening due to biochemical reactions was observed through a visible and near-infrared (VIS/NIR) spectroscopy (Muhammad and Peeyush, 2014). In their study, a portable VIS/NIR spectrometer is employed to rapidly measure quality of oil palm FFB on-site, by means of non-contact and non-destructive approach. Two statistical analyses are performed to models FFB quality, i.e., the ripeness, the OER and also the FFA. The former method, i.e., a forward-stepwise method is employed to establish multiple linear regressions. The latter method is a combination between principal component analyses with multilayer perceptron neural network (Muhammad and Peeyush, 2014). In conclusion, the measured color spectral data can be statistically analyzed to predict not only the OER, but also the FFA. Therefore, there should be relationships between the FFA with the color of FFB.

2.2. Digital image pre-processing

All images in this paper are taken by a digital camera with similar specs as the ones found on any smart phone. A digital image is defined as a discrete two-dimensional function, $f(x, y)$, where x and y are the spatial (plane) co-ordinates. The amplitude f at any pair of co-ordinates (x, y) is called the intensity (color) or gray level of the image at that point. A digital image is usually composed of a finite number of elements, each of which has a particular location and values. These elements are called "picture elements" or "pixels". The intensity (color) of each pixel is variable. Typically, a color is represented by three or four component intensities of a color space, for instance, the RGB color space is defined by the red, green, and blue additive primaries and the CMYK color space is defined by the cyan, magenta, yellow, and black colors. In our work, the RGB color space is used since it is a common choice for computer graphics due to its simplicity.