Journal of Food Engineering 119 (2013) 220–228

Contents lists available at SciVerse ScienceDirect

Journal of Food Engineering

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

Expert system based on computer vision to estimate the content of impurities in olive oil samples

P. Cano Marchal *, D. Martínez Gila, J. Gámez García, J. Gómez Ortega

Robotics, Automation and Computer Vision Group, University of Jaén, Spain

ARTICLE INFO

Article history:
Received 28 January 2013
Received in revised form 18 May 2013
Accepted 20 May 2013
Available online 1 June 2013

Keywords:
Olive oil
Olive oil analysis
Computer vision
Support vector machines
Artificial neural networks

ABSTRACT

The determination of the content of impurities is a very frequent analysis performed on virgin olive oil samples, but the official method is quite work-intensive, and it would be convenient to have an alternative approximate method to evaluate the performance of the impurity removal process. In this work we develop a system based on computer vision and pattern recognition to classify the content of impurities of the olive oil samples in three sets, indicative of the goodness of the separation process of olive oil after its extraction from the paste. Starting from the histograms of the channels of the *Red–Green–Blue* (*RGB*), *CIELAB* and *Hue-Saturation-Value* (*HSV*) color spaces, we construct an initial input parameter vector and perform a feature extraction previous to the classification. Several linear and non-linear feature extraction techniques were evaluated, and the classifiers used were Support Vector Machines (SVMs) and Artificial Neural Networks (ANNs). The best classification rate achieved was 87.66%, obtained using Kernel Principal Components Analysis (KPCA) and a grade-3-polynomial kernel SVM. The best result using ANNs was 82.38%, yielded by the use of Principal Component Analysis (PCA) with the Perceptron.

1. Introduction

The application of computer vision to develop tools to measure the quality of goods and the performance of processes is ubiquitous (Malamas et al., 2003), and the food industry is one of its main fields of application (Brosnan and Sun, 2002; Du and Sun, 2006; Ruiz-Altisent et al., 2010; Rosell and Sanz, 2012). However, there are not many applications of computer vision to the olive oil industry, being those mainly to detect defects in the olives (Diaz et al., 2004) and determine its maturity (Furferi et al., 2007, 2010).

The determination of the content of impurities in virgin olive oil is a very frequent analysis performed on virgin olive oil samples. It is useful whenever there is a bulk oil transaction, as the maximum content allowed of moisture and impurities is fixed, and if the olive oil contains higher levels than those, the price paid is decreased accordingly. Determination of impurities is also helpful during the olive oil elaboration process, as it enables the operator to check the performance of the moisture and impurities removal process held in the vertical centrifuge or the settling tanks. For this latter application of the analysis the required precision is lower, since it is used only to adjust the process. A sketch of the olive oil elaboration process is presented in Cano Marchal et al. (2011).

The official method established in the international norm ISO 663:2000 requires dissolving the olive oil sample with hexane and filtering it through a previously dried filtering paper. Afterwards, more hexane is filtered through this paper in order to remove any residue of oil, and the filtering paper is introduced into an oven to eliminate the hexane. Finally the hexane-free paper is weighed and the content of impurities determined. This method is quite work-intensive, and it would be convenient to have an alternative approximate method to evaluate the performance of the impurity removal process. To the best knowledge of the authors, results regarding the use of computer vision to estimate the content of impurities in olive oil samples are still scarce, and we have found no reference of any actual industrial application.

The goal of this research work is to develop an approximate method, less work-intensive than the official one, capable of discerning between different contents of impurities indicative of the wellness of the impurity removal process for its use as off-line feedback information by the operators of the process. For that objective we propose and develop a system based on computer vision and pattern recognition to classify the content of impurities of the olive oil samples in three sets.

1.1. Background and previous works

The application of computer vision to the olive oil elaboration process is still a fairly open field, since there are not many reported applications and, to the best knowledge of the authors, even less commercially available products used routinely in the sector.

* Corresponding author.
E-mail address: pcano@ujaen.es (P. Cano Marchal).

0260-8774/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.jfoodeng.2013.05.032

The majority of applications of computer vision to this field have been devoted to classify olives according to the presence of defects or to infer different properties of them, such as ripening index or oil content. Diaz et al. (2004) presented a comparison of three algorithms – Partial Least Squares (PLS), Mahalanobis distance and Artificial Neural Networks (ANNs) – to classify table olives according to their defects, and concluded that the best results are obtained using ANNs, with a classification accuracy of 90%. Riquelme et al. (2008) proposed using three consecutive discriminant analysis for the same objective, yielding a 75% of correct classification rate in the validation phase.

Regarding the ripening index of olives, Furferi et al. (2010) developed a system to predict it based on computer vision and a further refinement of the results using an ANN. This ANN combined the preliminary ripening index provided by the initial algorithms with chemical parameters obtained from historical curves from the region where the olives came from.

For the prediction of the oil content of olives, Ram et al. (2010) constructed two models based on linear regressions and ANNs. The best results were achieved using ANNs, obtaining linear correlations of 0.81 for Souri and 0.87 for Picual olives. To conclude with the computer vision applications to olives reported, it is noteworthy the work by Gatica et al. (2011), where they presented a method to recognise the diameter of the olives from images of the olive tree.

There are some works related to visual characteristics of olive oil that do not apply computer vision techniques, but that are worth mentioning. Moyano et al. (2008b,a) related the relationship between color parameters of the oils in different color spaces to their chlorophylls and carotenoids indexes using a spectrophotometer. In turn, Gordillo et al. (2011) studied the influence of turbidity grade on the color and appearance of virgin olive oil samples, using a turbidimeter and a spectrophotometer.

Concerning the applications of computer vision techniques to other liquid products in the food industry, it has been reported its usage to quantify the total quantity of bacteria in juice (Jin and Yin, 2010). In this work, the authors constructed a feature vector with different parameters, such as Area, Perimeter and Circularity, and used biomimetic pattern recognition to segment the bacteria and distinguish it from cells and impurities. Afterwards, a count of the segmented regions was held to determine the quantity of bacteria.

Lastly, it is relevant to highlight the work by Hepworth et al. (2004), where the authors employed computer vision to determine the size and velocity of bubbles in beer.

1.2. Problem description

The content of impurities is a determination that is usually demanded jointly with the determination of the moisture content of the sample, as these are substances exogenous to the olive oil and affect negatively its conservation. The method to determine the moisture content is to weigh a certain quantity of the sample and introduce it to a drying oven until the weigh of the sample is constant. Then, the moisture content can be calculated simply by means of the weigh difference of the sample before and after the drying process (Fig. 1).

During this drying process, the viscosity of the oil decreases as a result of the increase in the oil temperature – typically, the drying is held at 110 °C. This decrease of the viscosity allows the impurities to sediment in the porcelain evaporating basin. Fig. 1 shows an olive sample before it is introduced in the drying oven and after the drying process, where the deposition of the impurities is visible. This phenomena highlights visual differences between samples of different impurity content, thus suggesting the use of images of the samples after the drying process to construct an approximate method for rapid estimation of the impurity content.

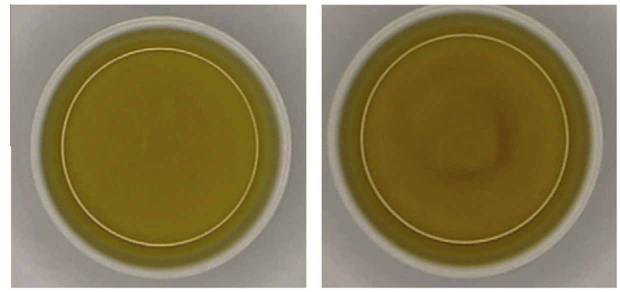

Fig. 1. Olive oil sample before (left) and after (right) drying process, illustrating the sedimentation of the impurities after the drying process.

The color of the olive oil samples is not constant and it is not related to the content of impurities. Besides, the patterns that the impurities form vary within the same level of impurities. Fig. 2 shows different samples for the three different levels of impurities considered, where the aforementioned differences between samples can be noticed.

The impurity content of the classes created for the classifier were chosen to be indicative of the goodness of the separation process, according to typical values indicated by practising olive oil elaboration experts. Impurity content below 0.04% may be regarded as low, a content between 0.04% and 0.05% (both included) is acceptable and values over 0.05% are indicative that the separation process can be improved.

2. Methods

The approach used was to select a candidate input parameter vector derived from the data contained in the images of the samples, then perform a feature extraction before feeding the data to the classifier. The following subsections detail each of these steps.

2.1. Experimental set-up

To get the measurement process constant during all the tests, different issues must be checked before acquiring the information. They are: no dust or dirt, low external light compared to our lighting system (reducing the noise coming from surround light), a constant temperature and no shock or vibration of the setup. All these recommendations are normally fulfilled in a chemical lab.

Fig. 3 shows the experimental set-up used for the image acquisition. The system was composed of a LZ836BP Logitech webcam of 2 MP resolution and a ring-shaped LED illumination device. The ring shaped lighting source used had a power of 17 W, a diameter of 18 cm, and was placed 25 cm above the samples.

The images were acquired using the software provided with the webcam with a Pentium IV standard desktop PC. The image processing, feature extraction and training of the classifiers were developed using an Intel Q9550 CPU, 4 GB of RAM, Windows 7 running PC. The software used to perform the experiments was Matlab, and the library used for the SVM implementation was LIBSVM (Chang and Lin, 2011).

The available sample set was composed of 154 samples, where 82 samples belong to the first class, 48 to the second and 24 to the third. The reason for such an unbalanced frequency of each of the samples is due to the typical occurrence of values which naturally occurs in the oil elaboration process, since the samples used were supplied by a laboratory specialized in olive oil analysis from the ones they received. An attempt was made to construct artificial samples for the higher level of impurities, but the samples