A Machine Vision system for real-time and automatic assessment of olives colour and surface defects

International Journal of Computer Research
Volume 17, Issue 3.
ISSN: 1535-6698

A MACHINE VISION SYSTEM FOR REAL-TIME AND AUTOMATIC ASSESSMENT OF OLIVES COLOUR AND SURFACE DEFECTS

Rocco Furferi [*] ***and Monica Carfagni***
Department of Mechanics and Industrial Tecnology
Via di Santa Marta 3, 50134, Firenze, Italy

ABSTRACT

The evolution of olive oil technology is related to research raised to deepen the understanding of biological and biophysical phenomena during the machining process olives, thus allowing the introduction of sensors able to monitoring the parameters and the process according to the characteristics of olives themselves. Current research has identified manufacturing protocols that can enable the achievement levels of product quality required by the market, mainly by reducing the time between collection and processing of olives and raising the technological level of extraction lines. In this context, the present paper aims to describe some of the results of a three-years project developed by the PIN-University of Florence (Italy) in collaboration with the Tuscan Regional Agricultural Development, the Florence Commerce Chamber "Laboratorio Chimico Merceologico—Azienda Speciale CCIAA di Firenze" and the SME "TEM" (Toscana Enologica Mori). The paper will describe the development of an innovative Machine Vision system able to 1) acquire, in real-time, images of olives in the conferring phase and 2) to process the acquired images in order to evaluate the ripeness of olives on the basis of their colour and the eventual presence of superficial defects. The devised system has been tested with the data extracted by olives in the harvesting period of year 2006.

1. PROBLEM FORMULATION

As wide known, two important parameters that affects the quality olives for olive oil extraction are the olives ripeness level and the olives sanitation condition. In literature [1, 2] has been demonstrated that these two parameters affect the quality of the olive oil independently from the process used for the extraction. If olives are characterized, for instance, by a high ripeness then the number of peroxides and the acidity level tend to increase and vice versa. The same occurs if the olives are defective. These parameters are

[*] E-mail: rocco.furferi@unifi.it

mostly important also because the oil mill works in strictly controlled condition. Accordingly the olive quality affects in a stronger manner the quality of extracted oil. The influence of the two cited parameters on the oil quality may be stated, averagely, in the range 40-50% and so these factors cannot be neglected when an high quality olive oil is the main aim of oil producers. A range of methods have been proposed for expressing the stage of maturity of olives. Among them the International Olive Oil Council has suggested a simple technique based on the assessment of the colour of the skins of 100 olives which are randomly drawn from 1 kg of a sample lot [3]. The first stage of ripening is known as the 'green stage', corresponding to green mature fruits that have reached their final size. After this stage the chlorophyll pigments in the olive skin are progressively replaced by anthocyanines during fruit ripening. This chemical process allows the possibility of identifying a 'spotted stage', a 'purple stage' and a 'black stage' according to the skin colour of the fruits [4]. Some experimental investigation was performed on virgin olive oils extracted from green, partially blackened and totally blackened olives in order to evaluate changes in the organoleptic properties and in the shelf-life of the oil. On the whole, the simple phenolic compounds increased as darker olives were used whereas the hydrolysable phenolic compounds decreased. In both the cultivars, total phenols and induction times were significantly higher in the oils obtained from green olives than in oils from totally blackened olives [5]. The visual control of the raw materials (i.e. olives) to be processed by the oil mills is a basic approach for automatically defining quality criteria about the typical product. The use of Machine-Vision (MV) based systems may help the human experts to perform a series of analysis to be, possibly, later processed. In Figure 1 the evolution of olive parameters (peroxides, polyphenols and oil content) and a comparison with the colour of olives is provided.

Accordingly the main objective of the present work is to describe a system for no-contact and real-time estimation of the colour and of the superficial defects of olives by means an image processing based approach.

The system has been developed according to the following tasks:

1. MV architecture definition.
2. Evaluation of the colour classes of olives by means of image processing algorithms.
3. Evaluation of the superficial defects

2. MV Architecture Definition

As wide known a Machine Vision system is mainly composed by an acquisition system, a proper illuminator and a series of dispositive for the link between the acquisition system and a PC. In the present work the acquisition system consists of a high resolution uEye UI-1480 camera QSXGA (2560x1920 pixel^2) provided with a ½ inches CMOS sensor and with a frame rate of 6 fps. The camera is rigidly attached to a support and positioned upright to the leaf remover – washing machine. The camera presents a spectral response both in the visible (RGB channels) and in the I.R. range as shown in Figure 5. In order to cut the response in the I.R. wavelength the camera is provided with a low-pass band IR cut filter at 650 nm. The camera is provided with an optic Tuss Vision LV0814 with Focal Length of 8mm, opening between 1 and 1.4 mm and angular openings equal to 56.5 ° (horizontal) and 43.9 ° (vertical) (see Figure 6). The camera is connected to a PC by means of a USB 2.0 connection, thus granting a maximum transfer rate of 480 Mbps. This transfer rate is sufficient for a 3-channels transmission of 6fps in full resolution.

The camera is provided with an optic Tuss Vision LV0814 with Focal Length of 8mm, opening between 1 and 1.4 mm and angular openings equal to 56.5 ° (horizontal) and 43.9 ° (vertical). The camera is connected to a PC by means of a USB 2.0 connection, thus granting a maximum transfer rate of 480 Mbps. This transfer rate is sufficient for a 3-channels transmission of 6fps in full resolution.

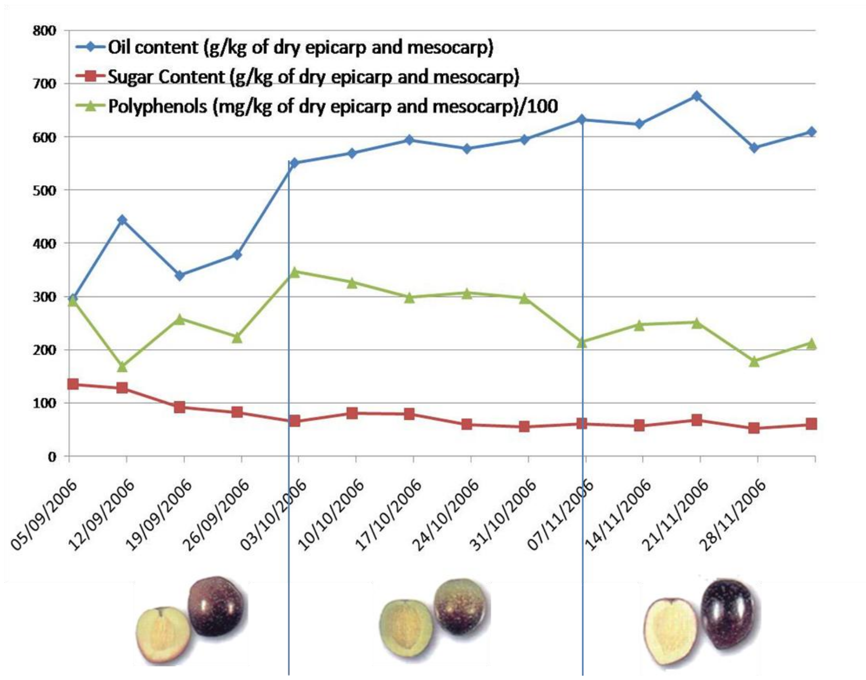

Figure 1. Evolution of oil parameters during harvesting.

The MV system is able, during the olive defoliating and washing, to acquire in full resolution with a frame rate of 6 fps. For this work it is sufficient to perform a quasi-static acquisition of the scene i.e.1 frame every 5 seconds. The images are acquired in RGB format. In Figure 2 the positioning of the camera upright to the leaf remover is showed.

As wide known by literature (and as, further, will be demonstrated below) the ripeness of olives depends by the colour of the olives themselves; therefore a proper illumination system should have to be used in order to perform a colorimetric analysis of each olive lot to be worked by the oil mill. Nonetheless the use of a black-box or of other systems for obtaining a strictly controlled illumination are not suitable for the purpose of the present work, being the camera attached upright the washing machine. For this reason the devised MV tool has to work in different light condition depending on the environmental circumstance and the colorimetric analysis has to take into account this problem. In any case an additional light consisting of a tungsten-halogen lamp with a temperature of 3623 K has been added to the MV system in order to illuminate the scene when the sunlight is lacking or missing.

Figure 2. Positioning of the CMOS uEye camera.

The images are then transmitted to a PC by means of a USB 2.0 cable. An ActiveX driver is used to set the camera settings; the driver is embedded into a Matlab® GUI.

In Figure 3 is shown an example of acquisition of an olive lot by means of the uEye camera.

Figure 3. Examples of acquisition of two different olive lots by means of the uEye camera.

For these reasons a scanner calibration target (used to calibrate all type of flatbed scanners or digital cameras) can be used in order to take into consideration any change in the light condition. For each frame the camera has to acquire both the olive lot into the washing machine and the calibration target (IT8 scanner target), as shown in Figure 4.

As wide known the IT8 scanner target is a printed piece that includes greyscale bars, and colour bars. Each value of the greyscale and each coloured bar, lead to a specific value in RGB or Lab colour space. By means of these values it is possible to create for each acquired image a proper correction by the following steps:

(i) The image of the olives with the scanner target within is acquired (in Figure 4 an image of an olive lot acquired in November 2007 of “Frantoio” cultivar is depicted).
(ii) The scanner target is isolated from the scene.
(iii) The R,G and B values of the columns 17, 18 and 19 of the target, that represents the standard values of red, green and blue (in different steps of brightness) are computed and compared with the R, G and B data stored by the manufacturer for the scanner target. The comparison lead to three values for each coloured area i.e. $\Delta R, \Delta G$ and ΔB, defined as the differences, in the three channels R,G and B, between the computed values and the stored ones.
(iv) The brightness values (L) of each gray scaled area are computed and compared with the brightness data stored by the manufacturer for the scanner target. The comparison allows the evaluation of a value ΔL for each gray scaled area, defined as the difference from the computed value of brightness and the stored one.
(v) $\Delta R, \Delta G$, ΔB and ΔL are used, image per image, to compute a colour calibration.

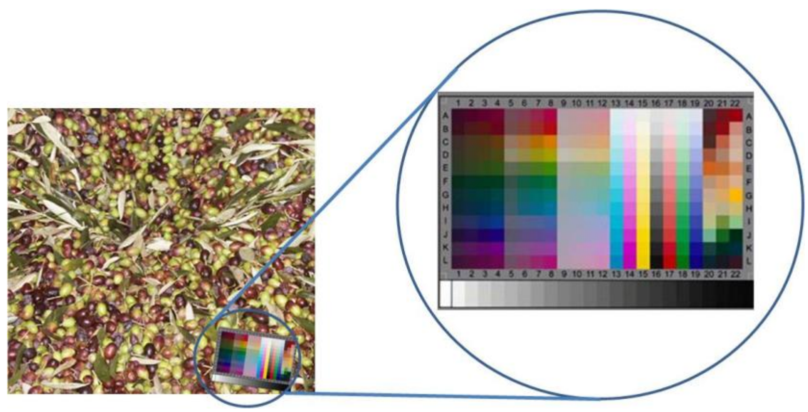

Figure 4. An example of acquisition of an olive lot by means of the uEye camera.

In this way for each image an ICC profile is created [6].

The MV system is able, during the olive defoliating and washing, to acquire in full resolution with a frame rate of 6 fps. For this work it is sufficient to perform a quasi-static acquisition of the scene i.e.1 frame every 5 seconds. The images are acquired in RGB format. Another task of the present work is the determination the sanitation conditions of each lot of olives. This condition depends on some factors, like for example the presence of olive fly (bactrocera oleae), that are not visible without a mechanical crush of the olive. Accordingly this factor can be related only to the presence of bruises, surface defects or advanced aging. The detection of this kind of defects in a image acquired in real time during the washing of olives is a very hard and time-expensive task. Accordingly, a number of olives are extracted from the lot and are manually disposed into a 20 cm x 20 cm grid for processing.

Figure 5. A selection of olives (lot 060042), positioned on a 20 cm x 20 cm grid.

In Figure 5 is shown a selection of olives (lot 060013), positioned on the described grid. Once disposed, the olives do not over impose each other and the risk of detecting false positive defects is minimized.

As previously mentioned, the described method for acquiring the images is suitable for the detection of the ripening grade and of the sanitation conditions of olives before the extraction process. Accordingly, after a description of the oil mill, will be described the image processing based algorithms devised by the authors for the assessment of this cited parameters.

3. Evaluation of the Colour Classes of Olives by Means of Image Processing Algorithms.

Once acquired and transferred to the PC, the images can be processed in order to extract a series of parameters correlated to the ripeness of olives. The main objective of the image processing algorithm developed for determining some parameters related to the ripening grade is to perform a colorimetric segmentation of each acquired image in order to detach the green olives in the image from the blackish-purple and brown olives. The devised algorithm performs a clustering of the coloured image into clusters characterized by different La*b* values. In other words, each acquired image is clustered into sub-images separated by colours. The image clustering allows to count up how many olives in the image are green and how many are blackish-purple and brown. The colour-based segmentation of the image may be performed by means of a K-means clustering based La*b* classification of objects acquired in the image. As wide known [7] the La*b* color space is derived from the CIE XYZ tristimulus values. The La*b* space consists of a luminosity L or brightness layer, chromaticity layer a* indicating where color falls along the red-green axis, and chromaticity layer b* indicating where the color falls along the blue-yellow axis. As a consequence an algorithm that performs a colour-based segmentation using the L*a*b* colour space [8] have been developed in order to count up the global area occupied by the green and the blackish-purple olives. The devised algorithm comprises the following steps:

1. Image colour conversion from RGB and La*b* colour spaces.
2. K-means clustering of La*b* image.
3. Detection of the area occupied by the differently coloured objects.

3.1. Image Colour Conversion from RGB and La*b* Colour Spaces

The first step for determining some parameters correlated with the ripeness of olives is to perform a conversion from RGB and LAB colour spaces. As known in literature, by means of the wide know eq. 3 it is possible to convert all the pixels values in RGB of the images into the tristimulus values CIE XYZ [9], under the illuminant D65:

$$[X,Y,Z]=[R,G,B]\cdot M_{D65} \quad (1)$$

where the matrix M_{D65} (size 3x3) is given by the Von Kries method [10].

The knowledge of the XYZ values, allows the colour transformation in the CIELAB space simply using the XYZ to CIELAB relations [11].

3.2. K-Means Clustering of La*b* Image

As wide known K-means clustering [12, 13] treats each object as having a location in space. It finds partitions such that objects within each cluster are as close to each other as possible, and as far from objects in other clusters as possible. K-means clustering requires the specification of the number of clusters to be partitioned and a distance metric to quantify how close two objects are to each other. In the present work a Squared Euclidean distance is used as metric. Since the colour information exists in the a*b* space, the objects are pixels with a* and b* values. In the present work a number of cluster equal to 5 is chosen in order to segment the green leaves, the white leaves, the green olives, the blackish-purple olives and the brown olives. In Figure 6 the results of La*b* colour conversion and of the following K-means clustering of an olive lot are shown. The original image is segmented into 5 images characterized by differently coloured object.

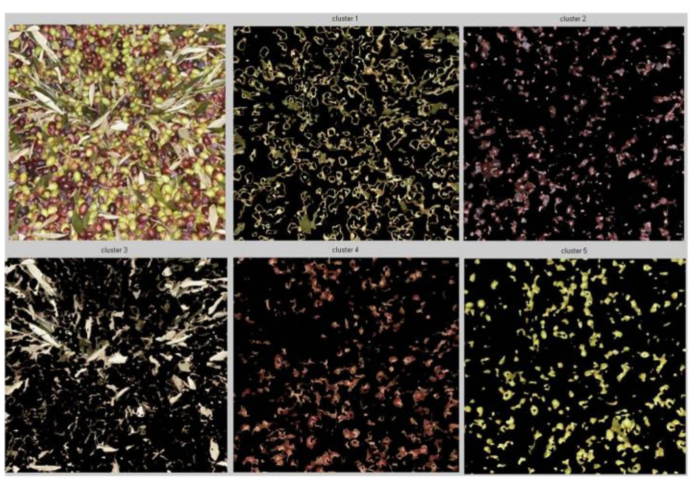

Figure 6. K-means clustering of a image of an olive lot.

Cluster 1 and 4 may be considered to refer to leaves. Cluster 2 shows the blackish-purple olives (advanced ripening) or the blackish-purple portion of olives with half ripening; cluster 4 the brown olives (optimal ripening) or the brown portion of olives with half ripening and, finally, cluster 5 shows the green olives.

3.3. Detection of the Area Occupied by the Differently Coloured Objects

Referring to clusters 2, 4 and 5 it is possible to estimate how many olives, in %, are in the considered lot by means of a simple formulation; the percentage of green olives G_O, for instance, is computed as the ratio between the area (in pixel) occupied, into the image, by the green objects in cluster 5 ($A_{clusters}$) and the global area of the acquired image (A_{image}):

$$G_O = \frac{A_{cluster5}}{A_{image}} \tag{1}$$

Analogously, it is possible to evaluate the percentage of blackish-purple (advanced ripening) and brown olives (optimal ripening); these parameters are called, respectively, B_{PO} and B_O are defined as follows:

$$\begin{cases} B_{Po} = \dfrac{A_{cluster2}}{A_{image}} \\ B_o = \dfrac{A_{cluster4}}{A_{image}} \end{cases} \tag{2}$$

The image processing task requires less than 5s to measure the olive ripening grade. This computational time is coherent with the fps settings described above.

In the example shown in Figure 22 the values of G_o, B_{Po}, and B_o are the followings:

$$\begin{cases} G_o = 15.54\% \\ B_o = 16.01\% \\ B_{Po} = 16.48\% \end{cases} \tag{3}$$

The three values G_o, B_{Po}, and B_o may be useful for an estimation of the ripeness of olives and will be used for software implementation; anyway it is not possible to affirm that the ripeness of olives is linearly dependent on the colour of olives (this is suitable only for that cultivars that changes their colour when mature). Unfortunately, a few varieties (cultivar) maintain a green colour even when ripen. Moreover a less mature green olive can be blackened either naturally (developed by maturity) or artificially (developed by oxidation). The natural blackening of olives is mainly attributed to formation and accumulation of flavanols, flavones and anthocyanines during maturation. Olives can be darkened by different methods, some of which utilize chemicals and even dyes in some locations [14] So, in order to assess an extensive definition of the ripeness of olives, it is required to understand the difference in green colour between the mature and the immature conditions and to assess if the olives have been treated artificially. Accordingly an experimental campaign has been conducted with the aim of test the devised algorithm in different conditions i.e. with different varieties of olives. In Table 1 the results of CIE La*b* conversion and K-means clustering are shown for 25 olive lots inspected during the experimental campaign. As described above, the devised algorithm allows a segmentation of each acquired image in 5 clusters where cluster 2, 4 and 5 are related to the olives.

Table 1. Results of CIE La*b* conversion and K-means clustering are shown for 40 olive lots inspected during the experimental campaign 2006

Olive Lot	Frantoio	Moraiolo	Leccino	Pendolino	Cluster 2 (B_{Po} %)	Cluster 4 (B_o %)	Cluster 5 (G_o %)
60009	100.00%	0.00%	0.00%	0.00%	6.66	29.85	21.19
60010	0.00%	0.00%	100.00%	0.00%	33.32	20.26	10.58
60011	0.00%	0.00%	0.00%	100.00%	5.97	8.41	43.65
60012	100.00%	0.00%	0.00%	0.00%	9.21	17.21	17.95
60013	100.00%	0.00%	0.00%	0.00%	9.57	14.8	39.03
60014	0.00%	100.00%	0.00%	0.00%	17.01	13.46	17.55
60015	100.00%	0.00%	0.00%	0.00%	9.75	17.11	28.05
60016	0.00%	0.00%	100.00%	0.00%	40.82	19.15	30.39
60020	100.00%	0.00%	0.00%	0.00%	21.84	24.30	13.94
60021	100.00%	0.00%	0.00%	0.00%	12.34	28.17	12.40
060022/1	100.00%	0.00%	0.00%	0.00%	6.40	20.17	19.62
060022/2	100.00%	0.00%	0.00%	0.00%	30.68	16.84	21.70
60024	0.00%	0.00%	0.00%	100.00%	30.37	29.55	6.97
60025	100.00%	0.00%	0.00%	0.00%	14.20	23.48	13.34
060038/1	0.00%	33.00%	33.00%	33.00%	17.456	18.81	18.12
060038/2	0.00%	33.00%	67.00%	0.00%	12.44	13.81	13.93
60039	0.00%	33.00%	33.00%	0.00%	49.23	9.13	15.72
60040	100.00%	0.00%	0.00%	0.00%	12.76	13.94	17.54
60041	33.00%	33.00%	33.00%	0.00%	58.25	14.31	9.08
060042/1	0.00%	0.00%	0.00%	0.00%	3.188	27.167	38.62
060042/2	0.00%	0.00%	0.00%	0.00%	17.98	40.50	11.15
060042/3	0.00%	0.00%	0.00%	0.00%	17.13	39.42	28.97
60049	33.00%	33.00%	0.00%	33.00%	51.86	15.96	12.47

Referring to olive lots labeled with code “60009”, “60014”, “60016” and “60024”, it is possible to check up the results of clustering for mono-cultivar olives, typical of Tuscany Region (Italy). The olive lot 60009 is composed by the cultivar called “Frantoio”. This very fruity, aromatic, and herbaceous cultivar becomes blackish-purple when ripen. Accordingly the MV system allows a correct detection of the ripening grade on the basis of the colour. The same occurs for lot 60014 (“Moraiolo”) that change its colour gradually during maturation.

A different case is represented by the lot 60024 (“Leccino”); this variety becomes blackish-purple early and simultaneously. Accordingly the clustering into colours green, brown and blackish-purple could not be suitable for a characterization of the ripening grade of olives (a low % of green olives will be present in the lot during the harvesting of olives). Fortunately the devised algorithm is able to cluster the images without information about the kind of colour to segment; it only requires a number of classes for classifying the coloured objects in the image. As shown in Figure 7 (a), by setting 5 classes for clustering for the lot 60024 it is possible again to segment the olives into three classes that are not green brown and blackish-purple but green, blackish-purple and black. Accordingly, without a loss of generalization, it is possible to state that the algorithm devised is always able to classify the olives on the basis of their colour and gives at least 3 clusters i.e. three classes of colours.

Now, what's happen when a multi-cultivar lot is inspected? The lot is composed by different varieties each one characterized by different evolution in ripening.

In this case the clustering may induce to some errors in classification of the ripening grade on the basis of colour. For instance it is possible to consider the lot 60049 that is composed by the three varieties Frantoio, Moraiolo and Leccino.

The results of clustering are shown in Figure 7 (b). The cluster 4 presents the olives whose colour is blackish-purple but probably these olives are partially belonging to the Frantoio variety (ripened) and mainly belonging to the Leccino variety with different grades of ripening. For this reason it is not possible to state in a reliable manner that the ripening grade of this olive lot is linearly correlated to the colour. This problem has been solved, nowadays, by correcting the values G_o, B_{Po}, and B_o with the percentage of varieties composing the olive lot (in this example the lot is composed by the same quantity of the three varieties). The correction of the parameters is made by means of the knowledge of which cultivar change colour gradually during ripening and which one change colour rapidly during maturation. The correction is made only in case of lots composed by more than one variety of olives.

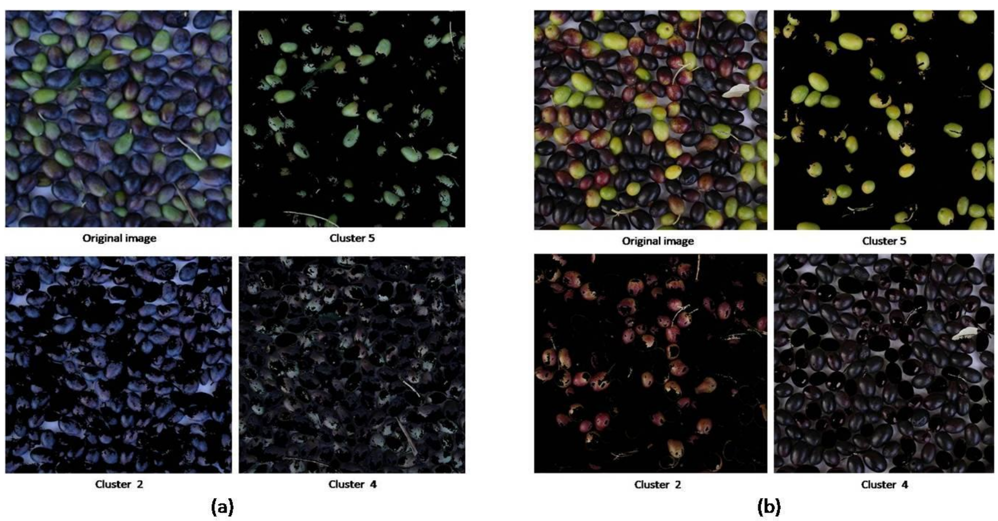

Figure 7. K-means clustering of an image of lot 60024(a) and of lot 60049 (b).

So, let:

– α_i be the percentage of the i^{TH} variety, composing the lot, that change its colour from green to brown to blackish-purple;

– β_i be the percentage of the j[TH] variety, composing the lot, that change its colour from blackish-purple to black i.e. presents a low % of green olives during the harvesting and change its colour rapidly during maturation. Obviously is:

$$\sum_i \beta_i = 100 - \sum_i \alpha_i \qquad (4)$$

The correction of the parameters G_o, B_{Po}, and B_o is given by the following equations:

$$G'_o = \frac{G_o}{\sum_i \alpha_i} \cdot (G_o + B_{Po} + B_o) \qquad (5)$$

$$B'_o = \frac{B_o}{\sum_i \alpha_i} \cdot (G_o + B_{Po} + B_o) \qquad (6)$$

$$B'_{Po} = \sum_i \alpha_i \cdot [(G_o + B_{Po} + B_o) - (G'_o + B'_o)] \qquad (7)$$

$$B''_{Po} = (G_o + B_{Po} + B_o) - (G'_o + B'_{Po} + B'_o) \qquad (8)$$

where G'_o is the corrected value for green olives and B_o is the corrected value for the brown olives. The blackish-purple coefficient B_o can be split in two parts: B'_{Po} is the percentage of blackish-purple olives belonging to the cultivars that change its colour from green to brown to blackish-purple and B''_{Po} is the percentage of cultivars that change its colour rapidly during maturation. In the case of lot 60049 the MV based algorithm B_{Po} = 51.85%, B_o = 15.96% and G_o = 12.47%; index i varies in the range 1-2 and index j is equal to 1. $\alpha_1 = \alpha_2 = \beta_1$ = 33%. Applying the equations described above it is possible to evaluate the following parameters:

$$G'_o = 15.17\%$$
$$B'_o = 19.41\%$$
$$B'_{Po} = 30.16\%$$
$$B''_{Po} = 15.53\%$$

Referring to the 40 lots inspected this situation lead to the evaluation of the above mentioned parameters for the lots 60049, 60052 and 60064. The results are listed in Table 2.

Table 2. Comparison between clustering results (G_o, B_{Po}, and B_o) and parameters G'_o, B'_{Po} and B''_{Po}

Olive Lot	B_{Po}'	B_{Po}''	B_o'	G_o'	B_{Po}	B_o	G_o
60049	30.16	15.53	19.41	15.17	51.86	15.96	12.47
60052	5.69	2.93	18.09	41.43	10.52	17.52	40.12
60064	28.29	9.43	7.31	6.81	31.41	10.57	9.85

4. EVALUATION OF THE SUPERFICIAL DEFECTS

In order to evaluate a parameter correlated to the sanitation condition of the olive lot, an image processing based approach has been developed. This approach is based on thresholding methods and performs the following tasks:

- RGB split into the three channels R,G and B.
- G-B subtraction and thresholding.
- Detection of defects on the green olives.
- Brown and blackish-purple olives defect detection
- Definition of a sanitation parameter

4.1. RGB Split into the Three Channels R, G and B

The acquired images are, numerically, composed by three matrices (array of numbers) of numbers varying in the range [0-255]. Each matrix represents the brightness values for the colours Red, Green and Blue. In Figure 8 is shown the effect of the split into the three channels R, G and B (i.e. in three arrays R, G and B) of the image of Figure 8; each channel allows the detection of different characteristics of the acquired image. In channel G it is evident the difference, in brightness, between the green olives and the blackish-purple ones. In channel B it is possible to see all the olives without distinction and it is evident the effect of reflexes of light on the olives.

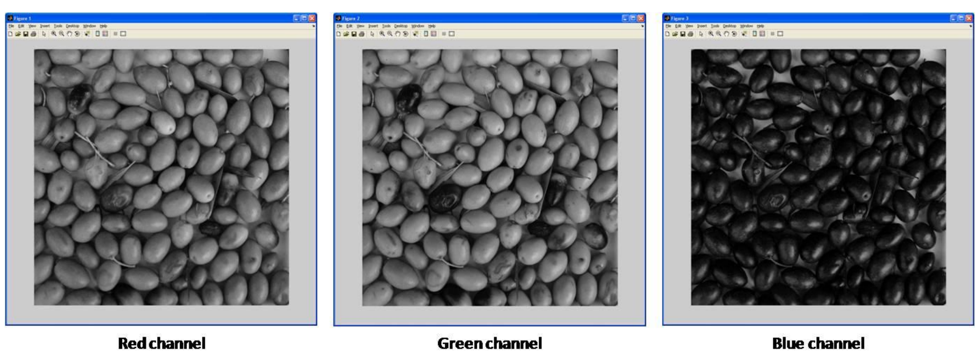

Figure 8. R, G and B channels for the image of Figure 24.

4.2. G-B Subtraction and Thresholding

Comparing the channel G with the channel B it is possible to see that black and blackish-purple areas do not change their brightness by changing channel. The same occurs for background. Accordingly if a each element of matrix B is subtracted from the corresponding element of array G it is possible to separate, in the image, only the green olives (with some leaves), as shown in Fig. 9. Let H be the array obtained as result of subtraction:

$$H = G - B \tag{9}$$

Now, the dark areas of olives in the array H represent both the defects of olives and the reflexes due to the light conditions. In order to detect only the defects on the green olives, a further segmentation of the image is required. For this reason a thresholding of image H can be performed in order to transform the dark areas of the image in white pixels. The thresholding is performed by means of a LTM method [15]. The result of this operation is a logical array BH (see Figure 10).

Figure 9. Image obtained subtracting B array from G array.

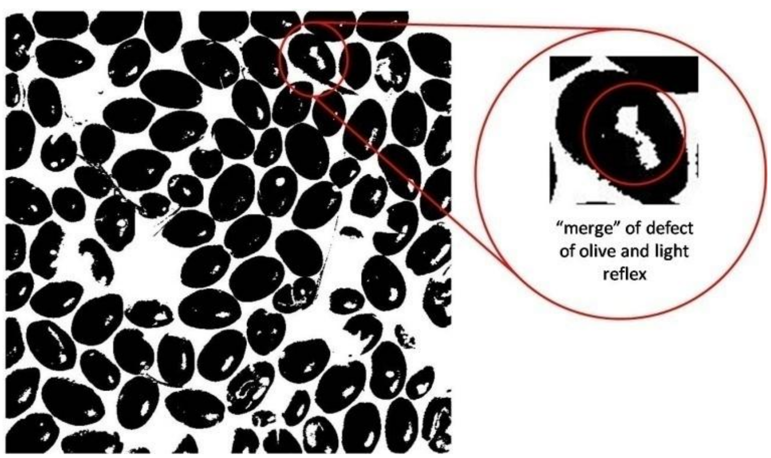

Figure 10. Thresholding of image H.

4.3. Detection of Defects on the Green Olives

In order to detect only the defects of green olives, it is suitable a further task: the thresholding of B channel. The effect of thresholding is to isolate, from the image, the contribute of light reflection on the olives, as shown in Figure 11. The result is a logical array BB. Now it is possible to detect only the defects on the olives by evaluating an array GD

(Green-Defects) obtained as the subtraction, element by element, of the two arrays BH and BB:

$$GD = BH - BB \tag{10}$$

In Figure 12 the comparison between the original image and the array GD shows the success in detection of defects on green olives (some defects are highlighted by a red circle).

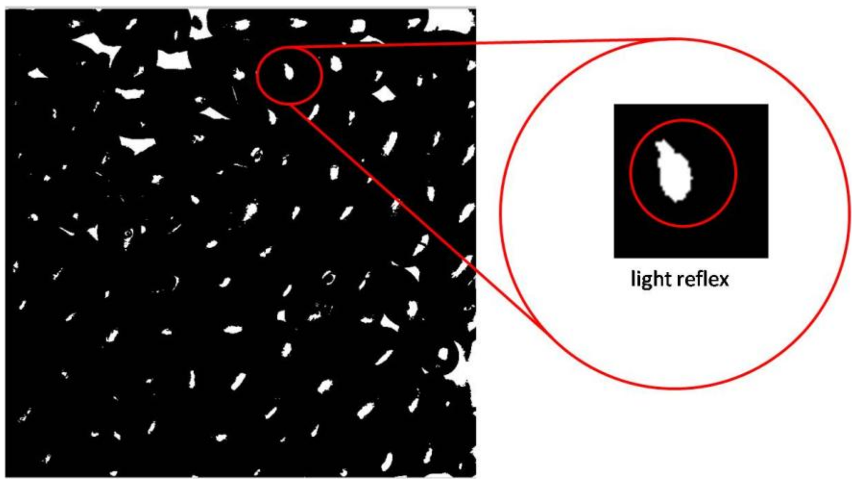

Figure 11. Thresholding of image B and detection of light reflex.

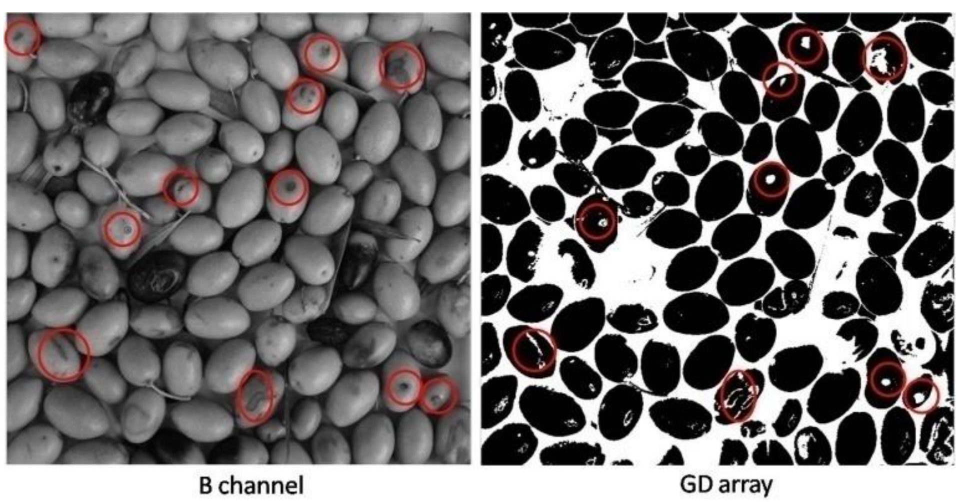

Figure 12. Detection of defects on the green olives.

Finally a blob analysis of the image GD allows the measurement of some geometrical properties of the defects and in particular of the area of defects for each green olive.

Another approach for detecting defects on the green olives is to perform a K-mean clustering of the image as described in the previous paragraph. According to this approach it is possible to segment the original image into three clusters, and consider only the cluster composed by the green olives. In this cluster the defects and the light reflex are represented by black areas on the olives surfaces. By performing a LTM based threshold of this cluster it is possible to create an array BH' very similar to the array BH. The detection of defects is then analogous to the approach described above (see Figure 13).

By experimental evidence (on the 40 olive lots inspected with the MV system) this approach, in the case of green olives, is less reliable than the approach described firstly.

Figure 13. K-means clustering based approach for defect detection on green olives.

4.4. Brown and blackish-purple olives defect detection

In order to evaluate the defects eventually occurring onto the brown, black and blackish-purple olives, a straightforward approach is to convert the grayscale image G to a binary image by means of a local thresholding. An example of this thresholding is given by the array BK obtained by thresholding the green channel of the image acquired for lot 060042. In Figure 14 this array is compared with the original RGB image.

Figure 14. Effect of local thresholding on the original image and definition of the array BK.

The dark areas on the surface of olives (in white) are resulting from the presence of defects and from the reflexes of light on the olives themselves. A separation of these two effects in case of dark olives is not a very simple task. Moreover the ripened olives are most subjected to different kind of defects as, for instance, bruising that not present an appreciable alteration of the brightness of olives. Accordingly the area of defects in blackish-purple, brown or black olives is approximately evaluated like the half of the area of dark spots in the Thresholded image of channel G with the supposition that the light condition is uniform for the inspected lot and that the reflexes influence these olives maximum for the 50% of brightness.

4.5. Definition of a Sanitation Parameter

The arrays GD and BK defined by means of the approach previously defined, allows the evaluation of some properties of the olive lot. More in detail by labeling the binary images and by using well known algorithms for blob analysis [16] it is possible to evaluate:

- The percentage of sane (undamaged) olives S_o id defined as the ratio between the area occupied (in pixel), in the image, by the undamaged olives (green, brown and blackish-purple) and the total area (in pixel) of the image. Referring to Figure 15, the green undamaged olives are represented by the dark areas highlighted with a yellow circle while the undamaged blackish purple olives are represented by the white areas highlighted by a red circle.
- The percentage of partially damaged olives P_D defined as the ratio between the area occupied (in pixel), in the image, by the partially damaged olives (green, brown and blackish-purple) and the total area (in pixel) of the image.
- The percentage of totally damaged olives T_D is defined as the difference between the total area A_O occupied by all the olives in the image and the two parameters previously defined follows:

$$T_D = A_o - P_D - S_o \tag{11}$$

Figure 15. Blackish-purple undamaged olives (red circles) and green undamaged olives (yellow circles).

Moreover it is possible to define a parameter called "Sanitation Parameter" as the ratio between the sum of all the areas occupied by the defects A_D (in pixel) and the sum of all the areas occupied by the olives A_O:

$$S_P = \frac{A_D}{A_O} \tag{12}$$

The parameters So, P_D, T_D and S_P will be used together with G_o, B_{Po}, B_o, B'_{Po} and B''_{Po} for defining an influence factor on oil peroxide number and on acidity and for training the ANN based software.

Conclusion

The present paper describes an image processing based approach for real-time and no-intrusive analysis of olive lots with the aim of 1) clustering the olives themselves on the basis of their colour and 2) to detect the superficial defects. The system allows a reliable clustering of images on the basis of their colour and this image segmentation may be useful for further development of the method. The defect detection will be further exploited in future works. Another task that will be developed in the future is addressed to the development of more effective image processing algorithms for the detection of superficial defects of olives, to the use of I.R. or U.V based image acquisition systems for the characterization of the olives and to the visual analysis of the olive oil extracted and filtered. In the present work the ripening grade of olives has been assessed by means of image processing methods on the basis of their colour. In a future development of the present work, the ripening grade will be measured also with the support of chemical analyses of the olives before the harvesting. At the present time the authors are working to devise a method for defining a more reliable Ripening Index (Jaen Index) by means of the combination of the proposed approach and a Neural Network based algorithm.

References

[1] Mugelli M., Cini E., Migliorini M., Simeani G., Modi G., Migliarese F., Sistemi di frangitura: influenza sulle caratteristiche chimiche ed organolettiche dell'olio di oliva extravergine, Firenze, IT: Ed. Camera di Commercio I.A.A. di Firenze Laboratorio Chimico Merceologico – Firenze, IT, 1998.

[2] Salvador M. D., Aranda F., Fregapane G. Influence of fruit ripening on 'Cornicabra' virgin olive oil quality A study of four successive crop seasons. *Food Chemistry*, 2001, Volume 73, Issue 1, pp. 45-53.

[3] Gómez-Rico A., Fregapane G., Desamparados Salvado M. Effect of cultivar and ripening on minor components in Spanish olive fruits and their corresponding virgin olive oils. *Food Research International*, 2008, Volume 41, Issue 4, pp. 433-440.

[4] Uceda M., Frías P. Épocas de recolección. Evolución del contenido graso del fruto y de la composición del aceite, Córdoba: Caja Provincial de Ahorros de Córdoba, 1975.

[5] Caponio F. , Gomes T. , Pasqualone A. Phenolic compounds in virgin olive oils: influence of the degree of olive ripeness on organoleptic characteristics and shelf-life. *European Food Research and Technology*, February 2001, Volume 212, Number 3. Springer Berlin / Heidelberg ,GE.

[6] T. Johnson. Methods for characterizing colour scanners and digital cameras, *Displays*, 1996, Vol. 16(4), pp. 83–191.

[7] Maier T.O., Rinehart C.E. Design criteria for an input color scanner evaluation test object, *J. Photo. Sci*., 1990, Vol. 38, pp. 169-172.

[8] Gonnet G.F. Colour effects of co-pigmentation of anthocyanines revisited-1. A colorimetric definition using the CIELAB scale, *Food Chemistry*, 1998, Volume 63, Issue 3, pp. 409-415.

[9] Kim D.H., Nobbs J.H. New weighting functions for the weighted CIELAB colour difference formula, *Proc Colour*, 1997, Vol. 97(1), pp. 446–449.

[10] Cini E., Daou M., Furferi R., Recchia L. A modelling approach to extra virgin olive oil extraction, *Journal of Agriculture Engineering*, 2007, Vol. 4 (1), pp. 1-10.

[11] Bordons C., Nunez-Reyes A. Model based predictive control of an olive oil mill, *Journal of Food Engineering*, January 2008, Vol. 84(1), pp. 1-11.

[12] Lloyd S. Last square quantization in PCM's, Bell Telephone Laboratories Paper, 1957. Published in journal much later: S. P. Lloyd. Least squares quantization in PCM. Special issue on quantization, *IEEE Trans. Inform. Theory*, 28:129–137, 1982.

[13] MacQueen J. B. Some Methods for classification and Analysis of Multivariate Observations, *Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley, University of California Press, 1:281-297.

[14] Kumral A., Basoglu F. Darkening methods used in olive processing, *ISHS Acta Horticulturae*, Vol. 791: V International Symposium on Olive Growing.

[15] Carfagni M., Furferi R., Governi L. A real-time machine-vision system for monitoring the textile raising process, *Computers in Industry*, 2005, Vol. 56, pp. 831–842.

[16] Horn B.K.P. (1986). *Robot Vision*. MIT Press, 69-71. ISBN 0-262-08159-8.