Optoelectronic Analyzer of Coffee Fruits

Juan R. Sanz-Uribe, Paula J. Ramos-Giraldo, and Carlos E. Oliveros-Tascón

Abstract— **This paper presents the development and implementation of a strategy to identify coffee fruits in five different maturation stages. The strategy was based in the separation of the points representing different maturation stages of coffee fruits, in the RG plane. To take advantage of this finding, a calibration algorithm complemented with the Convex Hull and Spline functions was developed to delimit the regions for the five maturation stages. An analysis based on expert criteria was utilized to make decisions in the intersections. The calibration procedure produces an identification matrix that is used to recognize the maturation stage of fruits passing in front of an optoelectronic device. The system showed promising results with a simple sensor and an inexpensive lighting system.**

Index Terms—**Coffee sorting, optoelectronics, identification matrix, maturation stages.**

I. Introduction

Coffee is the most important agricultural product in Colombia because millions of people depend on this crop. Colombian coffee is world famous for its good quality, which is due to a unique combination of conditions like soil, climate, agricultural practices, harvesting and processing.

Harvesting is carefully done by coffee pickers who detach and collect the ripe fruits one by one. However, the meticulousness that helps to obtain good quality also makes harvesting to account for the highest cost in coffee production; between 35 to 45%.

In order to confront the high harvesting costs, the Colombian Research Center of Coffee (Cenicafé) is currently working on developing technologies to increase the pickers´ capacity to levels at which they can increase their income and growers can reduce their production costs. Harvesting tools have demonstrated potential to drastically increase the picker's capacity; however, semi-mechanized harvesting is not as selective as manual harvesting. To take advantage of the economic benefit of the aided harvesting and to continue producing high quality coffee, there is a need for technologies to remove immature fruits, as well as undesired fruits and impurities, before processing.

Three processes are performed before obtaining ground coffee from the fruits. The first transformation is carried out in the farm and consists in the removal of the outer skin (pulping) of the fruits, removal of the mucilage that covers the seeds, classification and drying. This first process guarantees the coffee growers to sell a stable product that can be stored for months without problems of taste changes or decomposition by fungus. The second process is done by exporters and consists in removing the parchment-like husk of the grains and the third process is completed in the food companies and consists of roasting and grinding.

Coffee, like other agricultural products, is preferred by customers for its quality and consistency, but often the cup of the lot that arrives on port does not match the cup of the sample sent to promote the transaction. This is due to the distributed harvesting season, as fruits in all maturation stages are found in the trees during the year. The distributed harvesting makes the raw material inconsistent to process and justifies research on reliable technologies to maintain the same characteristics of the raw material during the year.

Another motivation for this research project is the market, which has been moving to coffees with special characteristics for which costumers pay significant extra money. A machine capable of separating coffee fruits in different maturation stages will trigger the possibility of exploring new flavors of mixtures of coffees with differing contents.

Physical properties, such as weight, shape, firmness, moisture and color are used to determine the maturity stage of fruits in destructive or nondestructive tests. Separating coffee in different maturation stages by optoelectronics is preferred because it is nondestructive and the necessary elements are reliable and inexpensive.

The purpose of this research work was to develop hardware based in optoelectronics and software to analyze and identify five different maturation stages of coffee fruits. The technology can also be used to identify different maturation stages of other fruits and agricultural products.

Manuscript received July 9, 2008. This work was supported by the Colombian Coffee Research Center (Cenicafé), the Colombian Science Foundation (Colciencias), and Promain Ltda.

J. R. Sanz-Uribe. Senior Researcher, Department of Agricultural Engineering, Cenicafé. Km 4 vía antigua Chinchiná-Manizales, Colombia. Phone: 57+(6)850-6550 (ext.403); fax: 57+(6)850-4723; e-mail: juanr.sanz@cafedecolombia.com.

P. J. Ramos-Giraldo. Associate Researcher, Department of Agricultural Engineering, Cenicafé. Km 4 vía antigua Chinchiná-Manizales, Colombia. Phone: 57+(6)850-6550 (ext.426); fax: 57+(6)850-4723; e-mail: paula.ramos@cafedecolombia.com.

C. E. Oliveros-Tascón. Principal Researcher, Department of Agricultural Engineering, Cenicafé. Km 4 vía antigua Chinchiná-Manizales, Colombia. Phone: 57+(6)850-6550 (ext.407); fax: 57+(6)850-4723; e-mail: carlos.oliveros@cafedecolombia.com.

II. Determining the Strategy

A. *Coffee Fruit Development*

After blooming, it takes approximately 32 weeks for the coffee fruits to reach maturity (Fig. 1). This maturation process can be divided into three stages, according to the changes experienced by the fruits [1]. The first stage, which runs until the 8th week, is characterized by a slow growth, the fruit is seedless, its color is shiny green and its shape resembles a match head. The second stage runs until the 16th week. During this time the fruit keeps its green color but volume and weight increase rapidly. After the 17th week seed formation begins and at the 26th week the outside color starts changing from green to red. The maximum development of the fruit is reached at the 32nd week, when the fruits have maximum matter content.

An average mature coffee fruit have an ellipsoidal shape with a length of 17 mm and is 14 mm wide and 13 mm high.

If the coffee fruits are left on the tree after maturation, the water and matter content decrease dramatically and the outside color becomes darker, which are characteristics of the post-maturation process.

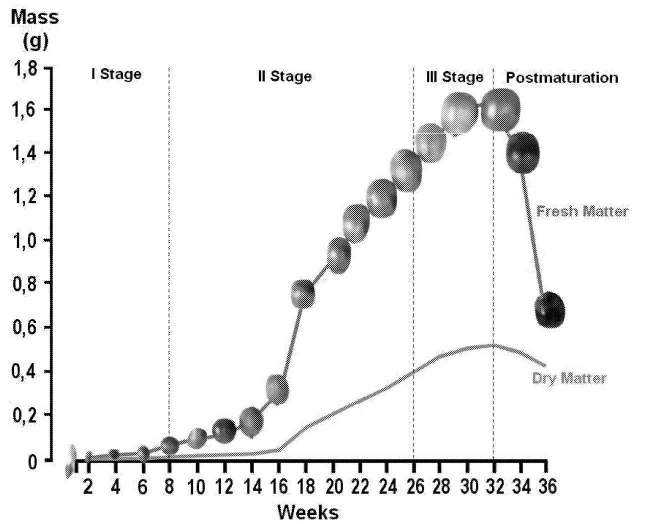

Figure 1. The development of coffee fruits

In this work, fruits are considered immature until the 26th week, undermature at the 26th week, mature at weeks 30 and 32, overmature at week 34 and dry after the 36th week.

The coffee quality is negatively affected by minimal contents of immature and dry fruits and, to a lesser extent, by considerable contents of undermature fruits [2], [3]. In general, mature and overmature fruits give good quality cups.

B. *Color Representations*

During the maturation process the outside color of the coffee fruit orderly passes in the spectrum of light from 551,5 nm, when immature, to 616,8 nm when mature [4]. This wide range is a characteristic that can be taken advantage of by optical means. Overmature fruits take a dark violet pigmentation, at about 450 nm in the spectrum, and dry fruits are brown.

Machine vision has been successfully used to sort different qualities in a large variety of agricultural products. Sometimes, binary images have been enough to create confident strategies to separate good grains from the undesirable ones, however, many times the information that can be extracted from these images is not sufficient.

There are several color representations that can be analyzed in order to create a confident strategy to separate the fruits in different classes. The most common color representation is the RGB, in which any color is described by its intensities of Red, Green and Blue. This color representation can be used with the three primary colors or some times one or two colors give the information needed to establish a confident strategy. Stogoff in [5] determined that, most of the time, the primary colors Red and Green give enough information to satisfactorily establish sorting strategies.

Another color representation commonly used to perform that task is using the descriptive parameters Hue, Saturation and Luminance, where hue describes the color, saturation the purity of that color and luminance its brightness. This color representation has shown advantages since is less affected by illumination changes [6]. The relationship between the RGB system and the hue, saturation and value (HSV) system is given by (1), (2) and (3).

$$v = \frac{r+g+b}{3} \qquad (1)$$

$$h = \begin{cases} \dfrac{1}{2\pi}\cos^{-1}\left[\dfrac{\frac{1}{2}\left[(r-g)+(r-b)\right]}{\left[(r-g)^2+(r-b)(g-b)\right]^{\frac{1}{2}}}\right], for\ {}^{b}\!/_{v} \le {}^{g}\!/_{v} \\ 1-\dfrac{1}{2\pi}\cos^{-1}\left[\dfrac{\frac{1}{2}\left[(r-g)+(r-b)\right]}{\left[(r-g)^2+(r-b)(g-b)\right]^{\frac{1}{2}}}\right], for\ {}^{b}\!/_{v} > {}^{g}\!/_{v} \end{cases} \qquad (2)$$

$$s = 1 - \frac{3}{r+g+b}\min\{r,g,b\} \qquad (3)$$

Where r, g and b are the intensities of Red, Green and Blue expressed in a decimal way, and h, s and v the normalized values of hue, saturation and luminance, respectively. The function min means the minimum value.

Another color representation that is being used more frequently is the one proposed by the *Comission Internacionale de l'Eclairage*, where all the colors can be expressed by a coordinate within a curved region of the *x-y* plane, which is called the Chromaticity Diagram [7]. This color representation is preferred because it is bi-variable and easier to work with, but it is not as sensitive as needed in this application.

C. *Setting up the Strategy*

In order to determine the identification strategy, five 200-fruit groups were prepared by experts with the maturation stages immature, undermature, mature, overmature and dry. The 1.000 fruits were photographed with a CCD camera of the brand Sony, reference Cybershot® DSC-P90, and a representative part of each fruit was extracted for analysis. The analysis was performed using the Vision tool of Matlab® and consisted in taking the intensities of red, green and blue of every sample. The data were plotted separately for each maturation stage in the planes RG, BR and BG planes, as well as converted to the HSV representation to observe for relevant differences in the hue axis.

Fig. 2 shows the RG plot, Fig. 3 the BR plot and Fig. 4 the BG plot. Fig. 5 shows the distributions of the different maturation stages in the hue axis. From these plots it can be seen that the RG plot presented the greatest potential to be used in the identification strategy because the clouds of points for every maturation stage are separated and well defined. From Figures 3 and 4 can be concluded that the color blue gives no maturation information.

The hue axis shows great potential to separate immature and undermature fruits because the distributions of these two stages are very concentrated, separated and with small intersection with each other. However, the intersection between undermature and mature fruits is significant, which leads to conclude a poor potential to be used as a base for the recognition strategy. Overmature and dry stages are very distributed in the hue axis and difficult to identify from each other and from the other stages.

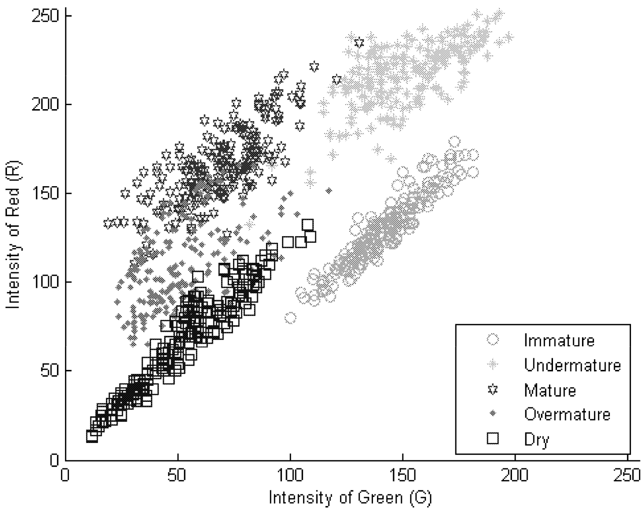

Figure 2. Plot of the five different maturation stages in the RG plane.

The proximity of the points representing fruits of the same maturation stage in the RG plot, leads to the opportunity of creating curved bounding lines enclosing all the points of each maturation stage. This was made by using the *convex hull* algorithm [8], which, with some modifications for this application, creates polygons with the most external points of every maturation stage, such as shown in Fig. 6.

The algorithm takes separately every maturation stage and records the four limit points: the minima values for R and G and the maxima values for R and G. The algorithm starts in the point with the minimum value for G, makes it the origin of a new coordinates system and obtains the slopes of all the points placed in its fourth quadrant, records the point with the smallest slope and converts it to a new origin. The procedure is repeated with the new origin until the point with the minimum intensity of red is reached. At that moment the procedure is repeated, but the slopes are calculated with respect to the points at the first quadrant, and until the point with the maximum intensity of green is reached. When the point with the maximum intensity of red is reached the slopes are calculated with respect to the third quadrant until the polygon is closed at the point with the minimum intensity of green. The same procedure is repeated with the other sets of points representing the other four maturation stages. The convex hull provides a two-column matrix with the most external points (vertices).

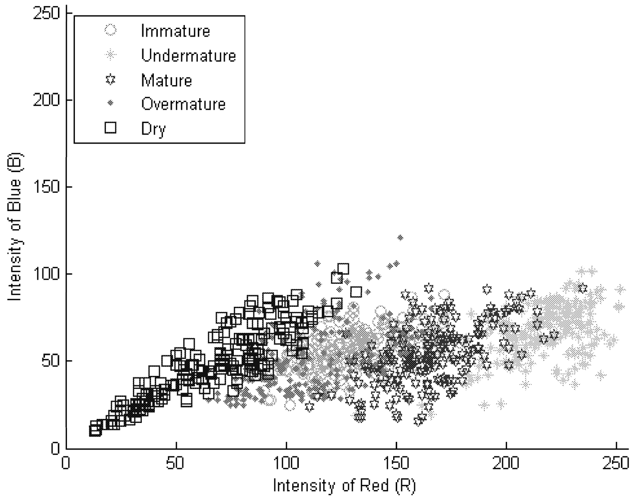

Figure 3. Plot of the five different maturation stages in the BR plane

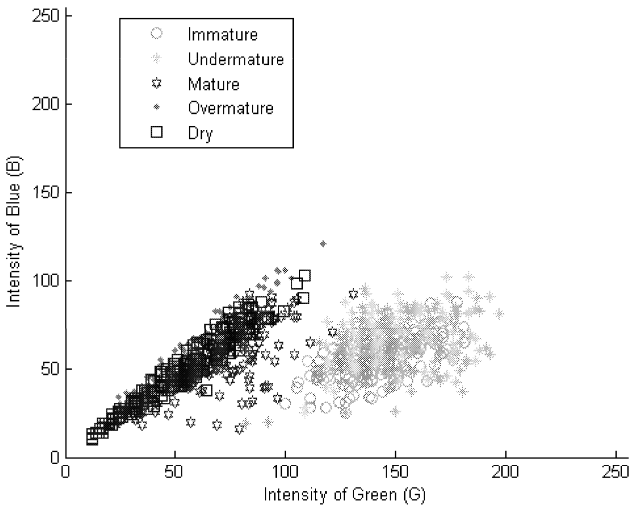

Figure 4. Plot of the five different maturation stages in the BG plane.

Once the five polygons are found, and the five matrices with the vertices are obtained, each of them are smoothed by using the *Cubic Spline Interpolation*, which is a procedure that introduces cubic polynomial curves (cubic Bézier curves) between each pair of vertices, meeting that two adjacent curves have the same first derivative and second derivative equal to zero at the common vertex, which make the joints smooth, relaxed and unnoticeable. With the purpose of the Bézier curve passing over the data points $\boldsymbol{S}$, it is needed to find a set of points $\boldsymbol{B}$, called control points, which serve to outline the curved line. The set of control points are found solving (4) for $\boldsymbol{B}$, with n the number of data points (vertices). In this way, the Bézier curves are influenced by the other points, making the interpolation part of a full smooth contour, such as shown in Fig. 7. In order to fix this methodology to the application, the starting vertex is taken in the part of the polygon where the slope has the minimum change.

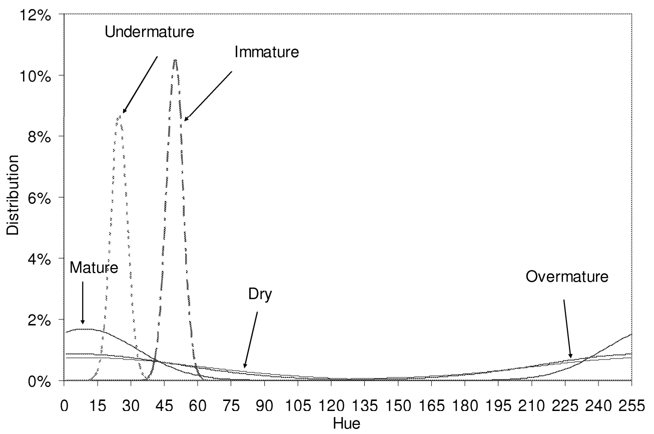

Figure 5. Distribution of the five different maturation stages in the hue axis.

$$\begin{bmatrix} 4 & 1 & 0 & \cdots & 0 \\ 1 & 4 & 1 & \cdots & 0 \\ 0 & 1 & 4 & \cdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & 1 \\ 0 & 0 & \cdots & 1 & 4 \end{bmatrix} \begin{bmatrix} B_1 \\ B_2 \\ B_3 \\ \vdots \\ B_{n-1} \end{bmatrix} = \begin{bmatrix} 6S_1 - S_0 \\ 6S_2 \\ 6S_3 \\ \vdots \\ 6S_{n-1} - S_n \end{bmatrix} \quad (4)$$

Fig. 7 also shows some regions common for two or more maturation stages. In order to know which maturation stage should be assigned to a fruit falling in an uncertain region, the percentage of occurrences, as well as the potential incidence on quality, were taken as criteria.

The intersecting region between mature and overmature contained the 34,5% of the total mature fruits, whereas the 25% of the overmature was included. This higher occurrence leads to a recommendation that all fruits falling in that intersection region are considered mature. The region common to overmature and dry fruits contained the 36,5% of the dry fruits and only the 10,0% of overmature fruits, leading to consider all the fruits falling in that uncertain region dry fruits. This last recommendation was reinforced by its positive influence in quality since dry fruits produce undesired flavors and tastes to the brewed coffee. The intersection section between mature and undermature fruits contained only the 2,5% of mature fruits and the 1,0% of undermature fruits, which leads to consider all the fruits falling in that area mature fruits. The common area among undermature, mature and overmature contained few fruits of which the greater amount was the 0,5% of the mature fruits, which permits all the fruits falling in that area to be considered mature. Finally, the intersection zone between overmature and undermature contained the 2,0% of the overmature fruits and only 0,5% of the undermature fruits, which allows all the fruits falling in that area to be considered overmature. Fig. 8 shows the regions after the mentioned analysis of occurrences.

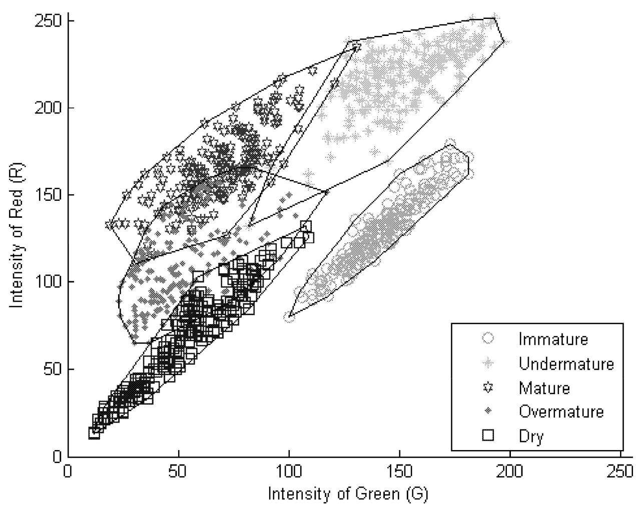

Figure 6. Polygons enclosing the five different maturation stages.

The next analysis consisted of creating an identification matrix containing all the information generated in the steps above. Fig. 8 shows the 256x256 identification matrix used to recognize five different maturation stages of coffee and elements that are not coffee fruits. In order to put the matrix in a digital way, the coordinates corresponding to the immature region were filled with ones, the undermature region was filled with twos, the mature region with threes, the overmature region with fours and the dry region with fives. The rest of the coordinates of the matrix were considered unknown and were filled with zeros. The identification matrix was evaluated with five 100-fruit groups corresponding to the five different maturation stages. The intensities of red and green of sample areas of every fruit, in the 0 to 255 format, were obtained and placed in the identification matrix. The maturation stage was assigned according to the region the fruit fall in. The evaluation gave the results shown in Table 1.

The high efficacy observed for this strategy validates it as a confident identification strategy to use in an autonomous electronic device with less complex sensors. The next chapter describes the implementation of this strategy with an inexpensive and simple sensor, since final cost should be as low that coffee growers can afford it.

III. EXPERIMENTAL SET-UP

The system has two functions: one for calibration and the other for identification. The function for calibration is when the system creates the identification matrix in the RG plane as shown previously, and the identifying function consists in recognizing the type of maturation stage that has every fruit that passes in front of the sensor.

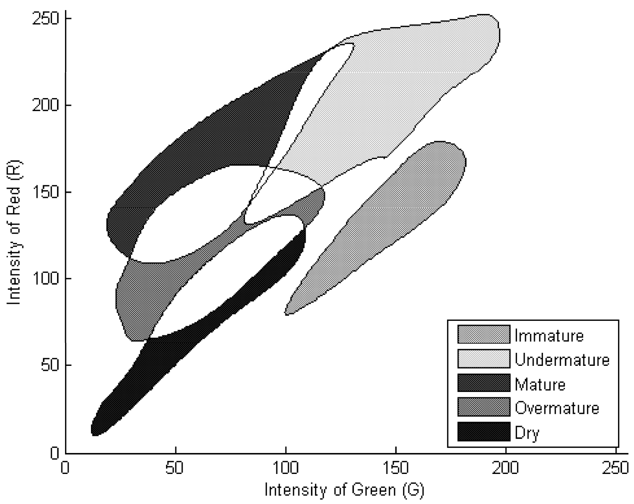

Figure 7. Smoothed polygons and intersection regions.

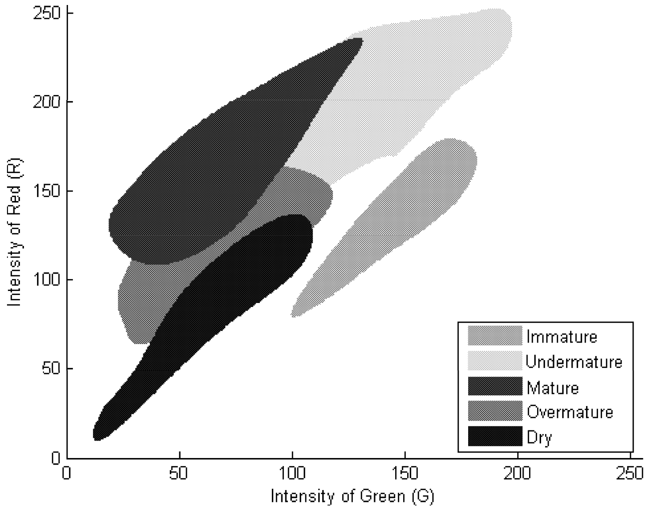

Figure 8. Identification matrix for five different maturation stages of coffee.

Table 1. Efficacy of identification

Identified / Actual	Immature	Undermature	Mature	Overmature	Dry
Immature	100%	0%	0%	0%	0%
Undermature	0%	97,5%	0%	2,5%	0%
Mature	0%	0%	100%	0%	0%
Overmature	0%	0%	10%	90%	0%
Dry	0%	0%	0%	0%	100%

The experimental set-up consisted of a desktop computer, a data acquisition electronic card of the brand National Instruments, reference DaqCard-6024E, a connection block CB-68LP of the same brand, a CCD sensor of the brand TAOS, reference TCS230, included in an AXE045 card.

The optical interface AXE045 has two white-light high-brightness LEDs for illumination and a 2mm lens that allows projecting a 1mm x 1mm image in a CCD of 64 pixels, of a 4mm x 4mm target placed at 25 mm from the lens. Each pixel has four photodiodes with red, green and blue filters and one with no filter for presence. The desired color to obtain the intensity is set by a two-bit protocol. The only output of the electronic device is a frequency proportional to the intensity of the measured color.

The experimental set-up also included a disk with 22 basins to carry the same number of fruits, driven by a stepping motor, to pass them in front of the optical interface. Fig. 9 shows the experimental set-up.

The commercial software used to program the algorithms for calibration and identification were MatLab® 11 and LabView® 7.0, respectively.

IV. Results

The algorithm for calibration described above was programmed in MatLab® and the hardware was set to acquiring color data in real time. Once the calibration program was run with five 100-fruit groups, corresponding to the five maturation stages, the resulting identification 256x256 RG matrix was loaded in memory to recognize the passing fruits. The identification matrix was loaded by means of LabView® in the desktop computer and set to identify fruits with the rest of the hardware.

Fig. 10 shows the identification matrix obtained with the mentioned sensor and lightning. It is noticed the illumination was not sufficient for this application because the data were grouped in the first eighth part of the graph (maximum 256 per axis), which means the environment was too dark for the sensor capacity.

Table 2 shows the results after running the identification function with other five 100-fruit sets, corresponding to the five maturation stages. The greatest efficacy was reached with mature coffee, followed very close by the one for overmature fruits. Surprisingly, the lowest efficacy was obtained with immature coffee fruits.

V. Conclusion and Recommendations

It was presented a strategy to identify coffee fruits in five different maturation stages. The strategy was based in an identification matrix that was created throughout a calibration procedure, which consisted in discomposing the fruit color in intensities of red and green and creating five regions in the RG plane, corresponding to immature, undermature, mature, overmature and dry fruits.

ISBN: 978-988-98671-0-2

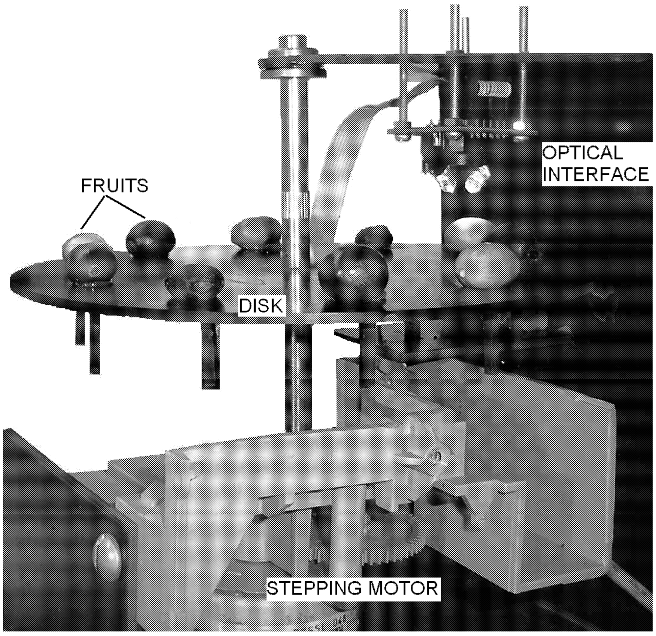

Figure 9. Experimental set-up.

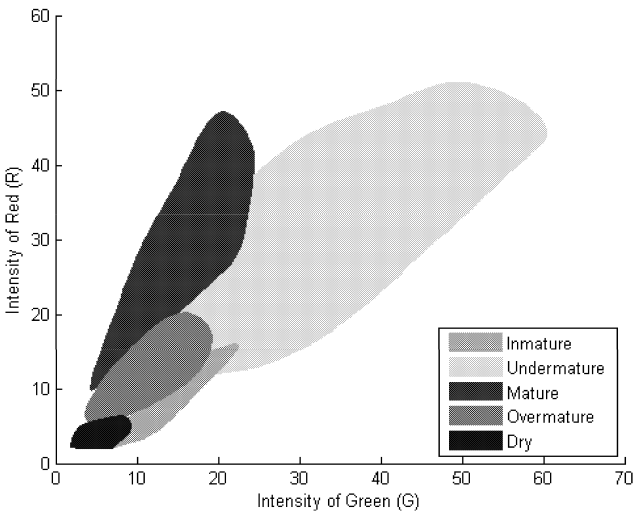

Figure 10. Identification matrix for the TAOS TCS230 sensor with two white-light high-brightness LEDs.

Even though there was a lack of illumination for the sensor capacity, the results are promising because it was an acceptable identification of mature and overmature fruits. It is recommendable to improve the illumination in order to widen the span of the identification matrix. The illumination can be improved by increasing the power and quality of the LEDs. Better performance is expected with a more powerful, stable and reliable illumination and the same optical interface.

The points that were considered unknown in the final evaluation ran very close to the regions. It is likely to perform an optimization study of the sample size to run the calibration process in order to obtain better sorting efficacies.

The occurrences analysis carried out to assign maturation stages to fruits falling in the uncertain regions is enough to make an acceptable decision. However, conducing neural networks and/or fuzzy logic analyses is recommended.

Table 2. Identification with the TAOS TCS230 sensor and two white-light high-brightness LEDs

Actual \ Identified	Immature	Undermature	Mature	Overmature	Dry	Unknown
Immature	60%	6%	0%	0%	6%	28%
Undermature	0%	74%	16%	8%	0%	2%
Mature	0%	0%	92%	4%	0%	4%
Overmature	0%	0%	6%	90%	4%	0%
Dry	4%	0%	0%	0%	64%	32%

It is also recommended to work in extracting the shades of the ends of the fruits due to the fruit curvature. These shades enter very dark color information that biases the calibration procedure.

The future work consists in bettering the illumination and implementing the strategy in an autonomous electronic device including the adaptive system that does not include the shades of the borders. A long term future work can be the implementation of the improved software and hardware in motorized manual machines to selectively detach coffee fruits from the trees.

REFERENCES

[1] M. R. Salazar, B. Chaves, N. M. Riaño, J. Arcila, Jaramillo R., A., "*Crecimiento del fruto de café /Coffea arabica/ var. Colombia*," Revista Cenicafé 45(2), pp. 41-50. 1994.

[2] G. I. Puerta, "*Influencia de los granos de café cosechados verdes, en la calidad física y organoléptica de la bebida*". Revista Cenicafé 51(2), pp. 136-150. 2000

[3] S. M. Marín, J. Arcila, E. C. Montoya, C. E. Oliveros, "*Caracterización de los Estados de Maduración del Fruto de Café*". Revista Cenicafé 54(3), pp.208-225. 2003.

[4] P. J. Ramos, C. Valdés, E. Gómez, J. R. Sanz-Uribe, and E. Solarte "*Características Espectrales de la Luz Reflejada por Frutos de Café (Coffea Arabica)*". Revista Colombiana de Física, Vol. 38, No. 2, 2006. pp. 822-825.

[5] P. E. Stogoff, *Sorting System*. U.S. Patent No. 1.926.824. Sept. 1933. [Online] Available http://patft.uspto.gov

[6] H. B. Nembhard, N. Ferrier, T. A. Osswald, and J. R. Sanz-Uribe. *"An Integral Model for Statistical and Vision Monitoring in Manufacturing Transitions"*. Quality and Reliability Engineering International (19) pp. 461-476. 2003.

[7] Y. Shirai, N. Kondo, and T. Fujiura. *Robotics for Bioproduction Systems.* ASAE 1998 ch.5. pp. 65-71.

[8] P. J. McKerrow, *Introduction to Robotics.* University of Wollongong. Australia. Addison-Wesley. 1990, pp. 454-467.