CONTROL OF THE COFFEE ROASTING STAGE USING ARTIFICIAL VISION TECHNIQUES

Juan Camilo Sarria-González[1], Eugenio Ivorra-Martínez[2], Joel Girón-Hernández[3]

ABSTRACT: Artificial vision techniques were used to evaluate its application in the control of the coffee roasting stage. Coffee samples of Colombia and Castillo varieties were obtained and analyzed by comparing images during the roasting stage. A one-way ANOVA analysis exhibited 94.28% of similarity of the coffee varieties studied; a multivariate analysis showed significant differences ($p<0.05$) for the time factor and its interaction with the variety factor, no differences were observed ($p>0.05$) for the coffee varieties. Additionally, a Principal Component, with two components demonstrated 90.77% of the variance by differentiating the samples in the different roasting times. Therefore, the proposed technique could be used in the control of the coffee roasting stage.

Index terms: Variety Colombia, variety Castillo, visible spectrum.

CONTROLE DO ESTÁGIO DE TORREFAÇÃO DE CAFÉ ATRAVÉS DE TÉCNICAS DE VISÃO ARTIFICIAL

RESUMO: Para avaliar o controle do estágio de torrefação de café, foram utilizadas técnicas de visão artificial e variedades de café, Colombia e Castillo, as quais foram analisadas através da comparação de imagens durante a torrefação. A análise de variância mostrou similaridade de 94,28% entre as variedades estudadas. A análise multivariada mostrou diferenças significativas ($p<0,05$) para o fator tempo e sua interação com o fator variedade, não foram observadas diferenças para as variedades de café ($p>0,05$). E ainda foi realizada uma análise de componentes principais. Com dois componentes principais, 94,23% da variância foi explicada pela discriminação das amostras nos tempos de torrefação. Ao que se conclui que a técnica proposta pode ser uma ferramenta no controle do estágio de torrefação de café.

Termos para indexação: Variedade Colombia, variedade Castillo, espectro visível.

1 INTRODUCTION

Since the late eighties, the introduction of cameras (visible spectrum) in different industrial processes has led to the emergence of control methods, based on the compression of the underlying information in an image using projection techniques on latent structures, known as Multivariate Image Analysis (MIA) (GELADI et al., 2006).

These techniques have been used for more than 20 years in agricultural applications related to robotics, automation, artificial vision and artificial intelligence (GARCÍA-LUNA et al., 2016). Furthermore, these techniques are quite powerful in situations in which images are used to record multitudes of color channels, collected by means of multispectral cameras, in which each channel can contain information of a smaller number of phenomena and therefore, be able to associate the demonstrated color channels with qualities and characteristics of interest to the study (PRATS-MONTALBÁN et al., 2011). Additionally, these techniques are often innovative and precise, and allow the development of technological and scientific advances.

Practical examples of its application in the control of food processes are: fat classification of Iberian pigs (GIRÓN et al., 2016), analysis in coffee roasting (HERNÁNDEZ et al., 2008), analysis in bread baking (VERDÚ et al., 2017), analysis of shelf life in fish (IVORRA et al., 2013; IVORRA et al., 2016) evaluation of pork meat quality (SUN et al., 2018), among others. Consequently, image analysis techniques demonstrate the effectiveness in predicting the behavior of a product under different conditions.

The expertise of a master roaster in the coffee roasting stage is still utilized, among other artisanal practices in the production of coffee in Colombia. This being said, it is necessary to identify low-cost practices and techniques that enable the processing of coffee to be optimized and controlled, thusly improving the efficiency of production and commercialization.

The analysis techniques that use artificial vision in the visible range provide access to information from which these processes can be

[1,3]Centro de investigación Cesurcafé - Universidad Surcolombiana - Neiva - Colombia - jcamilo.sarria@gmail.com, joel.giron@usco.edu.co
[2]Instituto de Investigación e Innovación en Bioingeniería - Universitat Politècnica de València Valencia - España - euivmar@upvnet.upv.es

automated and homogenized. The versatility of these techniques are applicable to the coffee roasting stage given that they allow for the quantification of data related to the changes in color and size that occur due to changes in temperature and time. The objective in this study was to develop and evaluate a procedure based on artificial vision for the control in the coffee roasting stage.

2 MATERIALS AND METHODS

20 kg of coffee (*Coffea arabica*) Castillo and Colombia varieties were harvested at an altitude of 1600 m.a.s.l., only beans with an intense red coloration were collected. The samples were deposited in polystyrene containers with cooling gels to avoid early fermentation and later were transported to CESURCAFÉ at the Universidad Surcolombiana (Neiva, Huila, Colombia) to perform the wet fermentation and drying process. In both varieties, the fermentation began with a pH value of 5.79 ± 0.013 until reaching a value of 4.56 ± 0.105. Subsequently, the drying process was carried out and the moisture content of the sample was reduced to $11.2\% \pm 0.04$ and $11.0\% \pm 0.06$ for the Colombia and Castillo varieties respectively.

'The hulling process was carried out in a laboratory hulling machine (ING-C-200, Ingesec, Bogotá, Colombia), followed by the selection of non-defective coffee beans in accordance with the provisions of the Specialty Coffee Association of America (SCAA, 2009). The coffee beans were passed through a number 13 sieve of 5 mm diameter to eliminate smaller beans. The roasting process was done in a Quantik roaster (TC-150R, Quantik, Colombia); two levels of roasting were defined, medium roast and dark roast, according to luminosity (L∗) measurements, similar to commercially available coffee samples (Craig et al., 2017), the roaster was set to sample entry temperature of 180 °C. For image acquisition, a Canon T3 digital camera (Tokyo, Japan) with 12 megapixels of resolution was adapted. The parameters of focal length, exposure time and white balance were set with an indirect lighting of 7.5 W dichroic LED lamps with 620 Lm (Lumek, Germany). To capture the images, Petri dishes containing the samples were placed 40 cm from the objective, 288 images were obtained in total, 6 images for each sample of 150 g of coffee and 24 captures for each time of roasting. For image analysis, it was developed an algorithm that segmented each captured raw image to eliminate background and noise (VALA et al., 2013) based on the parametric background subtraction algorithm (PICCARDI, 2004). From each element of interest (beans) a vector of morphological and color characteristics was extracted. Specifically, the average Red, Green, Blue (RGB) and Hue, Saturation, Value (HSV) coordinates values were obtained from a window around the centroid of each bean (LEE et al., 2017). In addition, the morphology was analyzed by textures of Haralick. Afterwards, information filtering with a conditional loop was performed by means of a dynamic thresholding to eliminate anomalous values (double the standard deviation). Finally, the data of each image was stored in a matrix to execute the statistical analysis. The processing of the images was done with the software MATLAB 8.0 (The Mathworks, USA).

Statistical analysis

'The similarities between images data of Colombia variety and Castillo variety were determined using a Student´s *t*-test. An analysis of variance (One-way ANOVA) was conducted in the first part of the study to determine whether there were significant differences. Each vector (color coordinates and morphology) was considered as a dependent variable and the variety was the factor. The least significant procedure (LSD) was used to test for the differences between means at a significance level of 5%. Additionally, in the second part of the study, a multifactor ANOVA was carried out for each vector to evaluate the effect of variety and roasting time and their interactions. In order to assess the feasibility of the artificial vision techniques to discriminate between coffee varieties and roasting times and to compress and select the information with the highest variance, a Principal Component Analysis (PCA) method was performed with a matrix of 44 variables obtained from the sensor, 30 for color (15 RGB and 15 HSV) and 14 for morphological features. All variance and multivariate analysis were carried out using Statgraphics centurion XVI program. (Manugistics Inc., Rockville, MD, USA).

3 RESULTS AND DISCUSSION

The similarity of the images obtained from the varieties studied presented a mean of 94.28%; the color coordinates showed a similarity of 93.33% for the RGB and 92.67 % for the HSV; a greater similarity was evidenced with the components of the shape vector 96.86%. The F-ratio and

P-values of the RGB coordinates obtained from the Colombia and Castillo varieties are shown in Table1, the values of the RGB coordinates only presented significant difference (p<0.05) for the maximum green variable. The constant variables in all the images were eliminated. Therefore, information of the minimum blue channel was not included in the analysis; ANOVA data for HVS and morphology information are not shown.

Figure 1 shows the interaction between the factors time and variety for a variable analyzed in this study (Hue mean) of the HVS coordinates. Differences were observed in the first minutes of roasting, and the effect of temperature decreased differences caused by the mucilage residues, similar results were reported by other authors (LEE et al., 2017); the response was not significant (p>0.05) for the varieties, while the roasting time values presented significant differences (p<0.05). This analysis showed similar results for all the parameters studied which justifies the application of the algorithm in different coffee varieties.

At each time, changes in color and shape of the beans were observed. In the roasting stage, these changes relate to the development of properties such as flavor, aroma and color (HERNÁNDEZ et al., 2008).

Figure 2 shows the scores plot in the plane of the first and second principal components (PCs) for the roasting stage, which together account for 94.22% of the total variability. Additionally, the value for each of the components established by the analysis is presented. From the third component, there are no significant increases for the variance, but 95.97% of the variability of the samples is noticed.

The analysis showed a good grouping of the samples in the study. The first component differentiated the coffee beans up to three minutes of roasting. In this time interval the beans still exhibited green colors and do not presented variation in volume. The second component separated the samples into four well-defined groups in which different sizes and color levels of roast were differed. At four minutes of roasting, when the internal temperature of the bean was 170 °C (Figure 3), the samples generated a group between the green color and the brown color and as the roasting process continued they got darker until reached minute seven when the temperature was approximately 220 °C (Figure 3).

At this time, the samples exhibited more homogeneous characteristics in both color RGB and HVS coordinates and shape morphology analyzed by textures of Haralick. Based on the results obtained, a robust classification algorithm capable of discriminating the roasting times can be performed.

TABLE 1 - F-ratio values and P-value obtained in a one-way ANOVA of the RGB color coordinates.

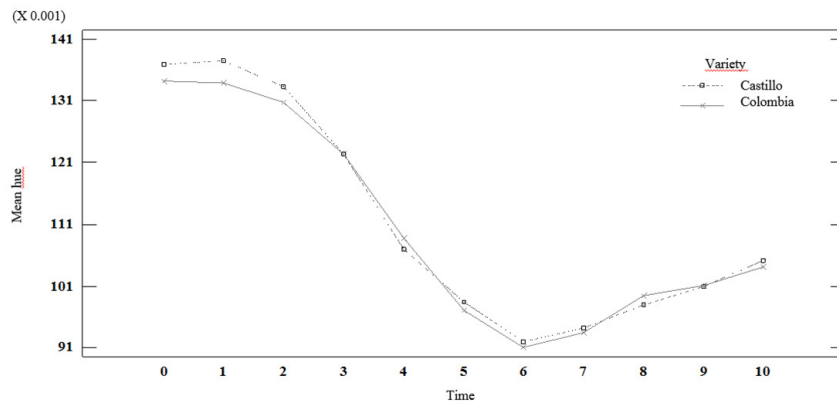| Variable	F-ratio	P-value
Mean Red	0.280	0.598
Mean Green	0.100	0.748
Mean Blue	0.050	0.816
Maximun Red	0.410	0.524
Maximun Green	3.950	0.048
Maximun Blue	1.340	0.248
Minimun Red	1.390	0.239
Minimun Green	0.060	0.806
Median Red	0.330	0.565
Median Green	0.150	0.698
Median Blue	0.060	0.814
SD Red	0.060	0.800
SD Green	0.180	0.674
SD Blue	0.260	0.612

FIGURE 1 - Interactions between factors coffee variety and roasting time for one of the evaluated variables, mean hue.
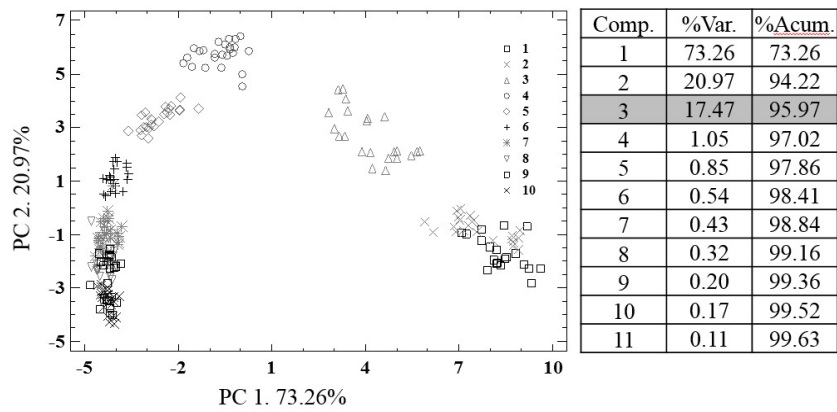

FIGURE 2 - Principal component analysis (PCA) performed on the image measurements for the roasting stage. With different times (1 (□). 2 (×). 3 (△). 4 (○). 5 (◇). 6 (+).

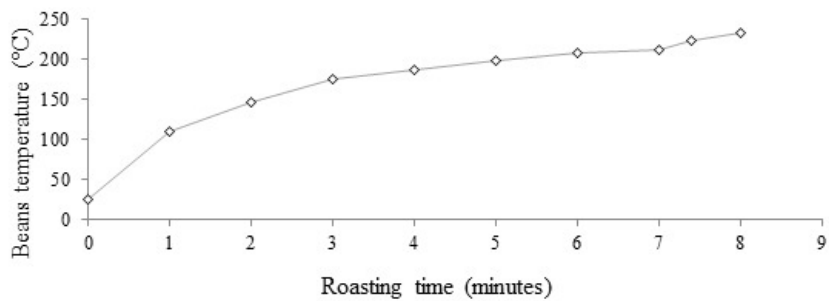7 (✳). 7.4 (▽). 8 (□) and 9 (✕) minutes).

FIGURE 3 - Roasting times and the temperature at which the beans came out of the roaster.

4 CONCLUSION

The application of artificial vision techniques allowed the automatic and objective differentiation of coffee samples according to the roasting time. The coffee varieties used in the trial were not a factor of discrimination. Future studies with other varieties are necessary in order to generate a classification tool applicable to any variety of coffee for the control of the roasting stage.

5 REFERENCES

Craig. A.; Botelho. B.; Oliveira. L.; Franca. A. Mid infrared spectroscopy and chemometrics as tools for the classification of roasted coffees by cup quality. **Food Chemistry**. V. 245. p 1052-1061. 2017.

GARCÍA-LUNA. F.; MORALES-DÍAZ. A. Towards an artificial vision-robotic system for tomato identification. **IFAC-PapersOnLine**, San Diego, v. 49, n. 16, p. 365-370, Oct. 2016.

GELADI. P.; GRAHN. H. Multivariate Image Analysis. **Encyclopedia of Analytical Chemistry**, New York, v. 1, n. 1, p.142-168 , Sept. 2006.

GIRÓN. J.; BARAT. J.; SÁNCHEZ. A.; GRAU. R. Aplicación de la espectrofotometría de infrarrojo (SW-NIR) para la clasificación de grasa de cerdos ibéricos. **Agronomía Colombiana**, Bogotá, v. 34, n. 2, p. 473-476, Nov. 2016.

HERNÁNDEZ. J.; HEYD. B.; TRYSTRAM. G. Prediction of brightness and surface area kinetics during coffee roasting. **Journal of Food Engineering**, San Diego, v. 89. n. 2, p. 156-163, Apr. 2008.

IVORRA. E.; GIRÓN. J.; SÁNCHEZ. A.; VERDÚ. S.; BARAT. J.; GRAU. R. Detection of expired vacuum-packed smoked salmon based on PLS-DA method using hyperspectral images. **Journal of Food Engineering**, San Diego, v. 117, n. 3, p. 342-349, Apr. 2013.

IVORRA. E.; SÁNCHEZ. A.; VERDÚ. S.; BARAT. J.; GRAU. R. Shelf life prediction of expired vacuum-packed chilled smoked salmon based on a KNN tissue segmentation method using hyperspectral images. **Journal of Food Engineering**, San Diego, v. 178, n. 4, p. 110-116, June 2016.

LEE. L.; TAY. G.; CHEONG. M.; BIN. P.; LIU. S. Modulation of the volatile and non-volatile profiles of coffee fermented with Yarrowia lipolytica: II. Roasted coffee. **LWT-Food Science and Technology**. v. 80, n. 1 p. 32-42, July 2017.

PICCARDI. M. "Background subtraction techniques: a review." IEEE International Conference on Systems. Man and Cybernetics (IEEE Cat. No.04CH37583), p. 3099-3104, Oct. 2004

PRATS-MONTALBÁN. J.; DE JUAN. A.; FERRER. A. Multivariate image analysis: A review with applications. **Chemometrics and Intelligent Laboratory Systems**, San Diego, v. 107, n. 1, pp 1-23,May. 2011.

Specialty coffee association SCAA. Grading Green Coffee. **Specialty Coffee Association of America**. USA: Santa Ana, Oct. 2009

SUN. X.; YOUNG. J.; LIU. J.; NEWMAN. D. Prediction of pork loin quality using online computer vision system and artificial intelligence model. **Meat Science**, San Diego, v. 140, n. 2, p. 72-77, June 2018.

VALA. H.; ASTHA. B. A review on Otsu image segmentation algorithm. **International Journal of Advanced Research in Computer Engineering and Technology**, Jawahar, v. 2, n. 2, p. 387-389, Feb. 2013.

VERDÚ. S.; VÁSQUEZ. F.; IVORRA. E.; SÁNCHEZ. A.; BARAT. J.; GRAU. R. Hyperspectral image control of the heat-treatment process of oat flour to model composite bread properties. **Journal of Food Engineering**, San Diego, v. 192, n. 1, p. 45-52, Jan. 2017.