International Conference of Agricultural Engineering
AgEng 2014 Zurich 6–10 July

Ref: C0343

Fraudulent barley detection in two different types of Robusta coffee by optical method

Evelin Várvölgyi, Lajos Dénes Dénes, János Soós, Zoltán Kovács, László Baranyai, József Felföldi, Department of Physics and Control, Corvinus University of Budapest, Somlói út 14-16., H-1118 Budapest

Gábor Szabó, Department of Process Engineering, Faculty of Engineering, University of Szeged, Moszkvai krt. 9., H-6724 Szeged

Abstract

The vulnerability to adulteration of coffee with regard to substitutions of physical characteristics (without indicating on the packaging) by several biological materials requires assurance of quality. In ground roasted coffee one of the typical frauds is roasted barley. In the work reported here a fast and low-cost optical method is proposed on one hand for the detection of barley as fraud in two different types of Robusta coffee. On the other hand the objective was the development of a method for prediction of the amount of fraudulent barley in coffee. Further aim was to develop a portable optical instrument to detect coffee adulteration in ground roasted form. The blending degrees of analyzed samples were between 0 and 50 wt% of barley in Robusta coffee. Moreover two commercial coffees with known barley-Robusta ratios were assessed. For vision system analysis we applied halogen and LED lighting as well. The HUE (angular location, 0 to 360°, of the color in the HSI system) spectrum of each sample (based on sum of saturation) was computed for statistical evaluation of vision system data. Six typical HUE values were selected based on their high correlation with the Robusta content of adulterated coffee samples. Based on PCA and LDA (calibration: 97.69%, validation: 96.55%) results LED lighting was more appropriate to discriminate and even to classify the samples by their Robusta content. Despite of the two types of Robusta the samples were overlapped containing the same amount of Robusta coffee. According to the results of PQS (Polar Qualification System) method the separation of the samples was not only due to the different amount of barley but the different types of Robusta as well along a quasi linear path. The amount of Robusta/barley in coffee samples was predicted by PLS regression method and by Radial SVM (Support Vector Machine) method. In case of LED lighting higher determination coefficient (R^2=0.996) and lower error were obtained (RMSEP=2.16%).

Keywords: coffee, barley, HUE spectra, adulteration, vision system

1. Introduction

After opening the coffee packaging the visual inspection is one of the first steps of quality control performed by the consumers. However impurities are difficult to detect in ground roasted coffee for naked eye if there is no reference. The most common frauds in coffee are cereals like barley, wheat, corn and soybean; and even brown sugar, husks and twigs (Assad et al., 2002). Nevertheless there are some countries where cereal-coffee mix is consumed, but if the mixing ratio is mislabeled the consumers are mislead.
Nowadays, digital image analysis is becoming more important because of its ability to perform fast and non-destructive, low-cost assessment on foods. Therefore it replaces the human visual system, often employed in quality analysis of food. CCD (Charge Coupled Device) cameras were applied to evaluate the external attributes such as the color of food. This type of cameras is able to create high-quality and low-noise images. Image pre-processing is a very important step of image evaluation (Du and Sun, 2004). Generally, color images are saved in the standard three-dimensional RGB (red, green, and blue) color space with that appropriate classification of pork and beef samples by fat/meat ratio and fat distribution was achieved by multivariate statistical methods (Felföldi et al., 2013). However HSI (HUE, Saturation, and Intensity) color system is more practical for color image distinguishing (Du and Sun, 2004).
The transformation of RGB color system to HSI was proved to be highly effective for the image processing of potatoes and apples in the research of Tao and co-workers (1995). In the inspection of color features of potatoes and apples represented by HUE histograms more than 90% accuracy was acquired by the vision system for discrimination between good and greened potatoes and yellow and green 'Golden Delicious' apples.
Abdullah and co-workers (2001) applied HSI color system to explore the relationship between oil palm fruit color and oil content for ripeness determination. The vision system classified correctly the oil palm at a greater than 90% success rate which is still higher than in case of human inspectors.
In another study, Du and Sun (2005) used HSI model by setting the HSI values in different ranges for segmentation of pizza sauce from pizza base and of light zones of pizza sauce from heavy zones.
Digital image and pattern recognition techniques have been used for edible vegetable oil classification applying RGB, HSI and grayscale systems in the work of Milanez and Pontes (2014). For expired and non-expired sample identification the HSI channels, while for classification of sunflower and soybean oils the HSI and gray channels were found to be more appropriate.
Sano and co-workers applied digital image analysis to quantify adulteration in roasted coffee (2003). Image processing of pictures prepared on Arabica coffee mixed with coffee husks and straw, maize, brown sugar and soybean was realized by gray-scale intensity scale. With this method good correlation was achieved (r=0.90-0.99) to quantify the amount of adulteration depending of different types of coffee substitutes.
Our objectives were to detect Robusta coffee adulteration by barley in ground roasted form by a vision system and to predict the amount of barley in coffee. Further aim was to develop a portable optical instrument to detect coffee adulteration in ground roasted form to help the work of authorities.

2. Materials and methods

2.1 Coffee samples and vision system setup

For the experiments two different types of commercial 100% Robusta coffee (indicated with K and P), a 100% ground roasted barley (B) and two commercial cereal coffees with known Robusta-barley ratios (AF-47R53B and Ot-51R49B) were received from a coffee factory.

Table 1 shows the prepared 24 admixtures with different amount of barley in Robusta coffee. Each sample weighted 18g.

Table 1: The Robusta content of Robusta (two types: K and P)-barley admixtures used for vision system analysis

No.	Robusta type K	Robusta type P	Barley content in wt%
1.	100RK	100RP	0
2.	99RK	99RP	1
3.	95RK	95RP	5
4.	90RK	90RP	10
5.	85RK	85RP	15
6.	80RK	80RP	20
7.	75RK	75RP	25
8.	70RK	70RP	30
9.	65RK	65RP	35
10.	60RK	60RP	40
11.	55RK	55RP	45
12.	50RK	50RP	50

The applied vision system was composed of a Hitachi HV-C20 3CCD with Canon TV Zoom lens. Image size was 768×576 pixel, and the optical resolution was 0.0833mm/pixel. The diffuse light was provided by special geometrical settings of 12 halogen lamps (20W, color temperature: 3200K) for the first experiment and by 6 SMD (Surface Mounted Device) LED panels (5.4W, color temperature: 5600K) for the second. For the experiment the sample was put in a circular, opened at the top, metal sample holder; the same quantity and uniform distribution of the sample was provided with the help of a ruler. After taking the picture, the sample was put back in its container and was mixed with the rest of the same sample for the next measurement. 10 repetitions of each sample were performed.

Images were stored in bitmap format (24 bit/pixel). RGB color parameters of each pixel of the image made by the camera were transformed into the HSI color system using the equations of Gonzalez and Woods (1992). HUE represents the angular location (0 to 360°) of the color in this system. Saturation values were collected into histogram according to the HUE angle (sum of saturation). Black and white components are not included in the evaluation because they do not contain color, in this way there is no need for segmentation algorithms contrary to other image processing techniques.

2.2 Statistical evaluation

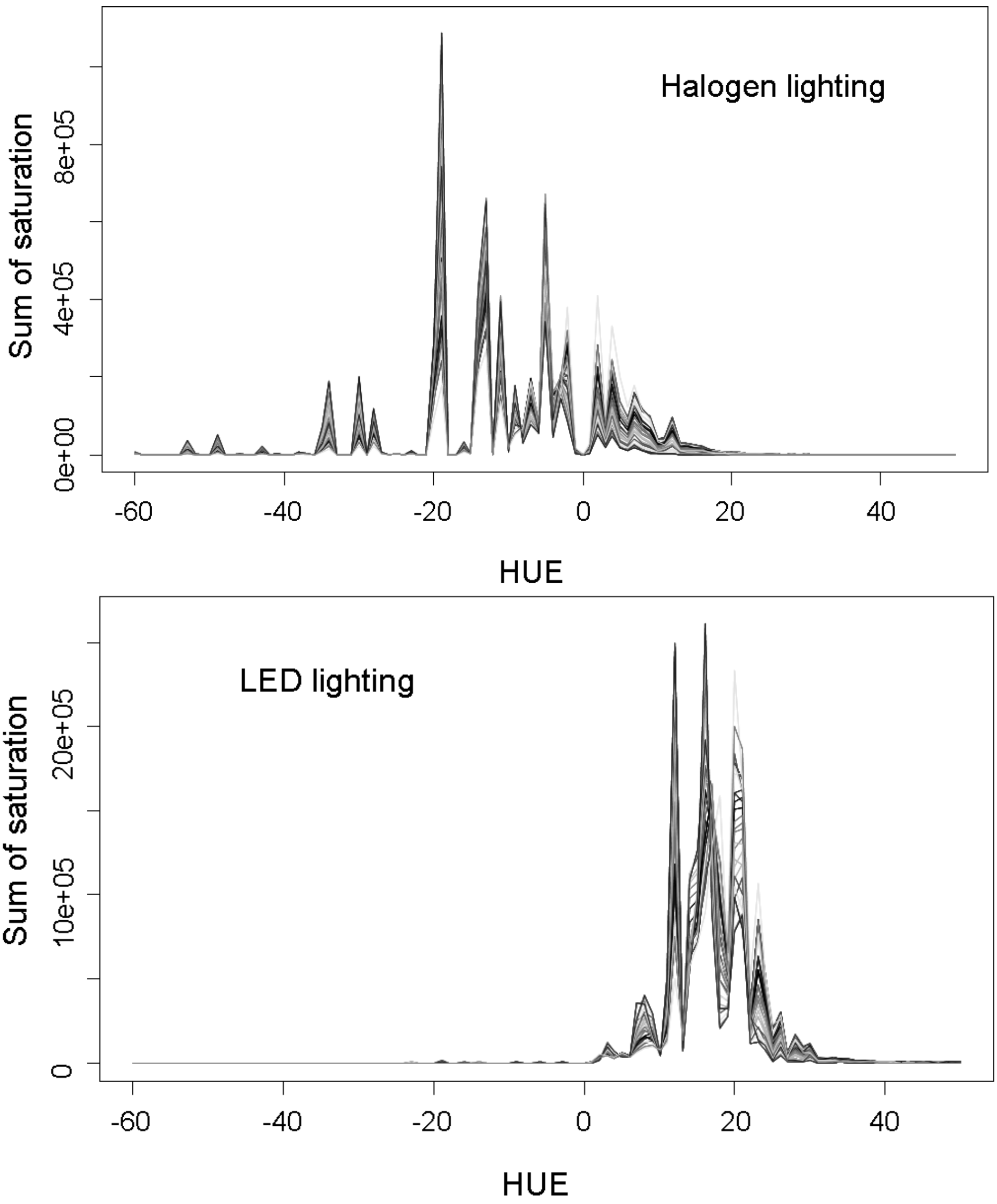
Statistical evaluation was realized by the software R 3.0.2 (R Foundation for Statistical Computing, Vienna, Austria). The statistical evaluation of vision system was performed using the non- supervised Principal Component Analysis (PCA) and the supervised Linear Discriminant Analysis (LDA) methods (Richards et al., 2002). Two third of the samples was used for calibration and one third for validation in case of LDA. Furthermore vision system data were assessed by Polar Qualification System (PQS) method. PQS is a general and powerful data reduction method rooted in the evaluation of NIR spectra. The quality of any spectra like data set is defined as the centre of its polar spectrum (polar coordinate system, where radius is the function of spectral value and angle is a function of wavelength. To compute coordinates of the quality point there are 3 approaches: the point, line and surface methods (Kaffka and Seregély, 2002). All mentioned approaches were applied for data evaluation.

Partial Least Square (PLS) regression and Radial Support Vector Machine (RSVM) was applied to predict the amount of barley in Robusta-barley mixture according to vision system results. All penalty parameters (gamma, epsilon and cost values) of SVM regression models were kept on default value. For SVM regression “radial” kernel and “eps-regression” were set. The obtained models were evaluated based on their determination coefficient (R^2), Root

Means Squared Error of Prediction (RMSEP) (Hill and Lewicki, 2007); Durbin-Watson (DW), Akaike Information Criterion (AIC) and Robust Parameter Design (RPD) values (Robinson et al., 2004). Akaike Information Criterion (AIC) is an indicator that is composed of the residual values and penalty derived from model parameters. Lower AIC indicates better fitting without overfitting. The validation was performed by LOO (leave-one-out) for PLS, in case of Radial SVM the same cross-validation was applied as for LDA.

3. Results and Discussions

Vision system raw data are presented on Figure 1, where the sum of saturation was represented in function of HUE values. In case of LED lighting the typical HUE range is more concentrated which is an advantage for data evaluation.



A

B

Figure 1: Typical HUE ranges of adulterated coffee samples in case of halogen (A) and LED (B) lighting

During the preliminary data evaluation typical HUE values were assigned according to the higher correlation with the Robusta content of adulterated coffee samples (halogen lighting: |r|>0.95, LED lighting: |r|>0.97). This way 6 different typical HUE values were selected containing relevant information in case of both lighting types.
According to the results of PCA the discrimination of adulterated coffee samples was successful by vision system. The first two Principal Components contained more than 99% of the data variance with both lighting types; however with LED lighting the discrimination of the samples was more efficient in the right order of increasing Robusta content. The discrimination of the two types of Robusta coffee was not significant.
LDA showed similar results than PCA, LED lighting was more appropriate to classify the samples by their Robusta content (calibration: 97.69%, validation: 96.55%). Despite of the two types of Robusta the samples were overlapped containing the same amount of Robusta coffee following the right order of Robusta content as it is shown on Figure 2.

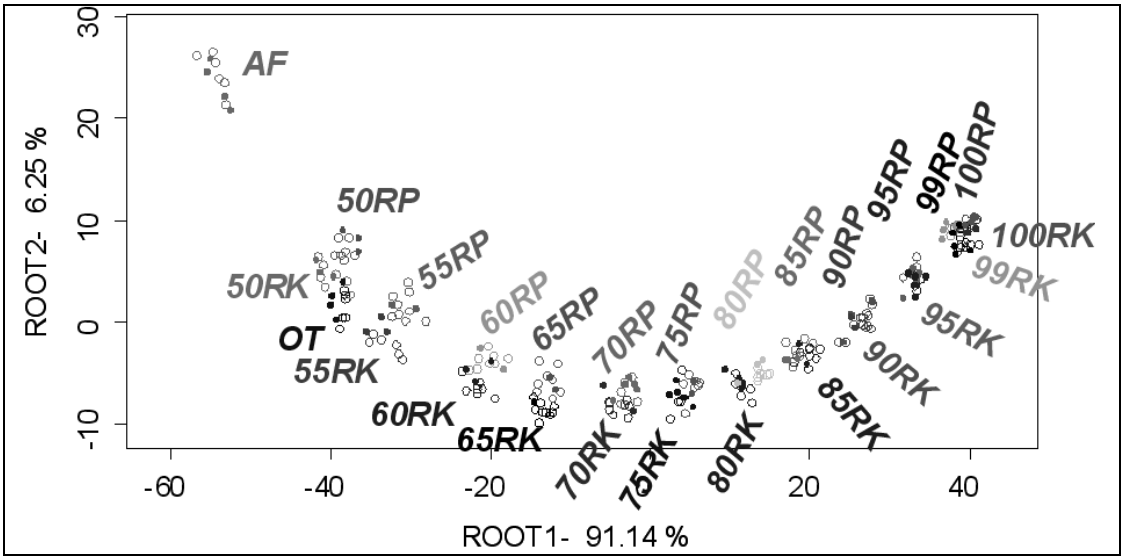


Figure 2: LDA of the adulterated coffee samples by barley in case of LED lighting, C: 97.69%, V: 96.55% (In sample names: P, K indicates the two types of Robusta, numbers show the Robusta content of the given sample-wt%, AF and Ot are commercial samples containing 47% and 51% of Robusta respectively)

Among the three approaches of PQS data reduction technique area method was found to be suitable in case of halogen lighting and point method for LED lighting. The represented points of coffee samples showed a monotonous change along a linear path in the right order of Robusta ratio. With the help of this data evaluation method the two types of Robusta were also discriminated.
The amount of Robusta in the coffee samples was predicted by PLS regression method. Table 2 shows that the coefficients of determination were significant with both lighting types; based on lower RMSEP value the LED lighting was found better. However DW values are not close to 2, therefore both models are loaded with autocorrelation.

Table 2: Model validation parameters of Robusta content prediction by PLSR in case of halogen and LED lighting

Model parameters	Halogen lighting	LED lighting
R^2	0.977	0.997
RMSEP, %	5.08	1.76
DW	0.84	1.05
AIC	231.24	1.03
RPD	6.54	18.51

While the linear model was proved to be inappropriate for prediction of Robusta content Radial SVM was performed. The parameters of established models are presented in Table 3; they are not loaded with autocorrelation. With LED lighting better results were obtained compared to halogen lighting.

Table 3: Model validation parameters of Robusta content prediction by Radial SVM in case of halogen and LED lighting

Model parameters	Halogen lighting	LED lighting
R^2	0.992	0.996
RMSEP, %	2.85	2.16
DW	1.953	1.737
AIC	31.539	10.356
RPD	11.41	16.13

4. Conclusions

Six HUE values were selected according to their correlation with Robusta content of the analyzed samples for data evaluation resulted by vision system. Based on PCA and LDA results the mixed samples with same amount of Robusta were not significantly discriminated by their Robusta type; however the sample groups followed the right order of Robusta content. Commercial samples (AF, Ot) were discriminated from the mixed samples. Using PQS method discrimination of the samples by the different type of Robusta was achieved following the right order of Robusta content. PLS regression was applied for prediction of Robusta ratio; however the established model was autocorrelated. Therefore Radial SVM was used for prediction of Robusta ratio in adulterated coffee mixtures providing appropriate results. LED lightening was found more suitable to detect coffee adulteration.
Compared to other method applied in the work of Sano and co-workers (2003) to detect adulterating materials in coffee, vision system does not require sample preparation, diffuse lighting minimizes the problem of shadows. Vision system can be suitable for adulterant barley detection in coffee even in low concentration, already in ground form to avoid mislabeling or fraudulence.

5. Acknowledgements

This research was supported by the European Union and the State of Hungary, co-financed by the European Social Fund in the framework of TÁMOP 4.2.4. A/2-11-1-2012-0001 'National Excellence Program'. Special thanks to Éva Vasné Hüvös for her professional advices.

6. References

Abdullah, M. Z., Guan, L. C., & Mohd Azemi, B. M. N. (2001). Stepwise Discriminant Analysis for colour grading of oil palm using Machine Vision System. *Food and Bioproducts Processing, 79 (4),* 223–231.

Assad, E. D., Sano, E. E., Correa, T. B. S., Rodrigues, H. R., & Cunha, S. A. R. (2002). Identificação de impurezas e misturas em pó de café por meio de comportamento espectral e análise de imagens digitais- Identifying adulteration in coffee powders by spectral signatures and digital analyses. *Pesquisa Agropecuária Brasileira, 37,* 211–216.

Du, C.-J., & Sun, D.-W. (2004). Recent developments in the applications of image processing techniques for food quality evaluation. *Trends in Food Science & Technology, 15,* 230–249.

Du, C.-J., & Sun, D.-W. (2005). Pizza sauce spread classification using colour vision and support vector machines. *Journal of Food Engineering, 66 (2),* 137–145.

Felfoldi, J., Baranyai, L., Firtha, F., Friedrich, L., & Balla, Cs. (2013). Image Processing Based Method for Characterization of the Fat/Meat Ratio and Fat Distribution of Pork and Beef Samples. *Progress in Agricultural Engineering Sciences, 9*, 27–53.

Gonzalez, R. C. & Woods, R. E. (1992). *Digital Image Processing,* Addison-Wesley, Reading, MA.

Hill, T., & Lewicki, P. (2007). *Statistics methods and applications.* StatSoft, Tulsa, OK.

Kaffka, K. J., & Seregély, Zs. (2002). PQS (polar qualification system) the new data reduction and product qualification method. *Acta Alimentaria, 31*, 3-20.

Milanez, K. D. T. M., & Pontes, M. J. C. (2014). Classification of edible vegetable oil using digital image and pattern recognition techniques. *Microchemical Journal, 113,* 10-16.

Richards, E., Bessant, C., & Saini, S. (2002). Multivariate Data Analysis in Electroanalytical Chemistry. *Electroanalysis, 14 (22),* 1533–1542.

Robinson, T. J., Borror, C. M., & Myers, R. H. (2004). Robust Parameter Design: A Review. *Quality and Reliability Engineering International, 20 (1),* 81–101.

Sano, E. E., Assad, E. D., & Cunha, S. A. R. (2003). Quantifying adulteration in roast coffee powders by digital image processing. *Journal of Food Quality, 26 (2),* 123-134.

Tao, Y., Heinemann, P. H., Varghese, Z., Morrow, C. T., & Sommer, H. J. III (1995). Machine vision for color inspection of potatoes and apples. *Transactions of the ASABE, 38 (5),* 1555-1561.