QUANTIFYING ADULTERATION IN ROAST COFFEE POWDERS BY DIGITAL IMAGE PROCESSING

EDSON E. SANO¹, EDUARDO D. ASSAD and SILVIA A.R. CUNHA

Embrapa Cerrados BR-020 Km 18 Cx. Postal 08.223 73.301-970 Planaltina, DF, BRAZIL

TANIA B.S. CORREA and HILDA R. RODRIGUES

Embrapa Agroindustria de Alimentos Av. das Americas, 29501 23020-470 Guaratiba, RJ, BRAZIL

Accepted for Publication March 18, 2002

ABSTRACT

A new method based on digital image processing used to quantify adulteration in roast and ground coffee beans is presented in this study. Pure arabica coffee and mixtures of coffee husks and straw, maize, brown sugar and soybean were produced in our laboratory as investigation materials. Red/Green/ Blue (RGB) color composites, magnified twelve times, were generated using a Charge Coupled Device (CCD) camera connected to a stereo microscope and a personal computer with an image processing software package. The percent areas of the contaminants in each image were calculated by the Maximum Likelihood supervised classification technique. Best-fit equations relating weight percentage (g.kg⁻¹) and the percent areas were obtained for each coffee contaminant. To test the method, 247 coffee samples of different amounts and types of adulterants were analyzed in the laboratory. The results showed that the new method developed can analyze precisely and quickly a large number of ground coffee powders.

INTRODUCTION

The detection of fraudulent or accidental adulteration of foods is an important issue for food manufacturers, regulatory authorities and retailers because of possible negative economic and health consequences for consumers.

Journal of Food Quality 26 (2003) 123–134. All Rights Reserved. ©Copyright 2003 by Food & Nutrition Press, Inc., Trumbull, Connecticut.

¹ Contact author. Edson Eyji Sano, Embrapa Cerrados, BR-020 KM 18 CX. POSTAL 08223, 73301-970 Planaltina, DF, Brazil. EMAIL: sano@cpac.embrapa.br

Therefore, suitable laboratory methods to identify and quantify the mislabeling of food samples are required by both food industries and government agencies to monitor the quality of food production.

There are three groups of adulteration in commercial roast and ground coffees (Briandet *et al.* 1996): adulteration with coffee substitutes such as barley, brown sugar, caramels, coffee husks and straw, maize, soybean, etc.; mixing of two coffee species, for instance, cheaper *robusta* (*Coffea canephora* Pierre ex Froehner) with pure *arabica* (*Coffea arabica* Linnaeus) coffee; and mixing expensive coffee beans from one growing region with cheap beans grown in another region. This study deals only with the first type of adulteration.

Regarding the first group, Brazilian coffee specialists classify coffee husks and straw as *impurities* because their presence in coffee powders, in most cases, are not intentional. They become part of coffee during the harvesting and husking stages. On the other hand, the presence of substances such as brown sugar, caramel, maize and soybean are considered intentional and are therefore classified as *mixtures*. Coffee adulteration is particularly important in Brazil because this country is the world's largest producer, with approximately 25 million bags per year (21 million of *arabica* and four million of *conillon*, the Brazilian variety of *robusta*) and since 1995 (Knight 1997) is the world's second most important consumer.

In Brazil, the first method to detect and quantify adulteration by coffee husks in roast powders was proposed by Menezes Jr. and Bicudo (1958). The method was based on visual microscopic analysis of coffee powders previously treated with chloroform. The coffee husks were separated manually through the use of a stereo microscope and quantified with a standard area diagram. An evolution of this technique is the use of an analytical scale instead of an area diagram (Instituto Adolfo Lutz 1985). Both methods are costly and time-consuming because they do not use advanced computer-based data processing. The goal of this study was to present a new operational and low cost method based on a digital image processing technique to quantify contaminants in roast coffee powders.

MATERIALS AND METHODS

Coffee Samples

As investigation materials, pure, roast coffee powders as well as samples adulterated with coffee husks and straw, maize, brown sugar, and soybean were prepared under laboratory conditions by the Brazilian Association of Coffee Industries (ABIC). The following concentrations of contamination were prepared: 100, 200, 300, 400 and 500 g/kg (g.kg⁻¹). Specifically for coffee

husks and straw, we also included contaminations of 10, 20, 30, 40 and 50 $g.kg^{-1}$ because of the regulatory authorities who allow the presence of this adulterant up to 10 $g.kg^{-1}$ (Brasil 1999).

ABIC also supplied 247 samples of laboratory-controlled roast and ground coffee powders with different amounts and types of contaminants. Some samples also presented two different adulterants in order to evaluate the performance of the method when more than one contaminant is introduced in the coffee powders. The envelope containing ABIC's results was opened only after the analyses were completed. All coffee samples presented a high degree of roasting, which represents more than 90% of all commercial coffee in Brazil.

Removal of Surface Oils and Fine Particles

Approximately three grams of homogenized, roast and ground coffee from each sample were transferred to Erlenmeyer flasks and allowed to mix overnight in 90 mL of chloroform in order to remove the surface oils and to destroy the aggregates present in the original samples. The mixture was then filtered using filter papers (fine porosity, slow flow rate and particle retention of $3 \mu m$). Fine particles (less than 0.297 mm of diameter) were also removed by using a 50-mesh sieve. Such a removal was necessary due to the difficulty in determining, quickly and precisely, whether a grain of less than 0.3 mm in diameter is coffee or a contaminant. The remaining materials were stored at room temperature in 4-cm diameter Petri dishes.

Image Processing

The basis of this study is that different materials present distinct spectral signatures. This is a well-known fact for the remote sensing community, in which different vegetation, soil, geological and land use classes, among others, have been identified in the optical range of electromagnetic spectrum (e.g., Madeira Netto 1996; Sabins 1997). Specifically in this study, we assume that the spectral reflectance of pure coffee is different from its substitutes in the visible range (400 to 700 nm). To test this assumption, representative RGB (Red/Green/Blue) color composites with 500 g kg⁻¹ of brown sugar, coffee husks and straw, maize and soybean were selected in order to compare digital numbers (DNs) of the adulterants with those of pure coffee. As shadows were present in all images, we also included them in the analysis.

Regarding DNs, any satellite or microscopic image consists of a certain number of small and equal areas known as picture elements (or pixels) arranged in regular rows and columns. Each pixel has a numerical value called digital number, digital count or brightness value. DNs range from zero to some higher number (e.g., 64 in a six-bit system or 255 in an eight-bit system, which is the case of this study) on a gray-scale intensity scale. Pixels with low numbers are related with materials having low reflectances and vice-versa. The number of bits available is determined by the design of the system (sensitivity of the sensor). The RGB color composites are obtained when three images of the same material, obtained in different wavelengths, are displayed at the same time, using the three channels (red, green and blue) available in the image processing software package.

The average DNs and the corresponding standard deviations for contaminants were obtained using the PC-based SPRING, the Brazilian image processing software package developed by the National Institute for Space Research (Inpe) (Camara *et al.* 1996). Using the cursor available in the SPRING, we selected representative training areas with varying number of pixels (8 - 624 pixels) for each contaminant. It is important to highlight that any other commercial or public domain image processing software package such as ENVI, IDRISI or ERDAS/IMAGINE, among others, could be used to process the microscopic images.

The samples were magnified 12 times with a stereo microscope. A visual inspection through this microscope is conducted in order to determine if the sample is pure or not. If not, a RGB (Red/Green/Blue) image is produced of the coffee sample with a CCD (Charge Coupled Device) camera connected to the microscope. The size of the images was 480×620 pixels and they were saved in a Tagged Image File Format (TIFF). A total of 15 images were analyzed for each proportion of contamination: three sample replicates and five images per replicate.

The images were classified with the Maximum Likelihood supervised classification technique available in the SPRING. In this type of classification, the analyst defines small areas in the image called training sites for each spectral class. Pixels located in these training sites form the training samples used to guide the classification algorithm in assigning all remaining pixels in the image to a specific spectral category. The Maximum Likelihood classification uses the mean and standard deviation values of each class, obtained from the training sites, to classify the entire image. More details of this technique can be found in Campbell (1987).

One of the statistical results given by the supervised classification is the percentage of the area occupied by each spectral category in a given RGB image. This means that, in this study, for a specific proportion of contamination by weight (g.kg⁻¹), there are 15 results obtained by the SPRING. The average of these 15 values was compared to its corresponding value in weight. Therefore, the final result of the digital image classification is the determining of the relative areas occupied by a given contaminant (coffee husks and straw, maize, brown sugar or soybean) in the following concentrations of contamination by weight: 100, 200, 300, 400 and 500 g.kg⁻¹. For coffee husks and straw, we also included five other proportions of contamination (10, 20, 30, 40 and

50 g.kg⁻¹), as reported previously. Finally, the best-fit equations relating these two variables (concentration of contamination and relative area occupied in the image) were obtained for each adulterant using the MicrosoftTM Excel software package.

RESULTS AND DISCUSSION

Spectral Separability Between Pure Coffee and Adulterants

Typical RGB color composites of both pure coffee and coffee contaminated with coffee husks (500 g kg⁻¹) are shown in Fig. 1a and 1b, respectively. In the pure coffee image, the dark colored spectral class corresponds to the shadows of coffee powders, while the yellowish class corresponds to the coffee powders themselves. Coffee husks and straw appear in the image with grayish and bright yellowish colors, respectively.

Straw presented the highest DN in all bands (237, 234 and 166 in the red, green and blue bands, respectively), while the shadow, as expected, presented the lowest values (81, 84 and 79 in the red, green and blue bands, respectively). The coffee powders and coffee husks presented intermediate DNs.

The separability of straw, pure coffee, coffee husks, shadows, brown sugar, maize and soybean in the Red, Green and Blue bands is also shown in graphic form in Fig. 2. Although in the blue band we can notice some confusion among the contaminants, mainly due to higher standard deviations, the separability is clearer in the other two bands.

Best-fit Equations

Figures 3 and 4 show the relationship between the concentration of adulterants and their relative area in the RGB image for coffee husks and straw, maize, brown sugar and soybean. The following equations were obtained:

Coffee husks & straw:	$y = 15.54 \ln(x) - 30.33$	$r^2 = 0.99$	(1)
Maize:	y = 0.06 x + 0.26	$r^2 = 0.93$	(2)
Soybean:	$y = 11.17 \ln(x) - 44.13$	$r^2 = 0.94$	(3)
Brown sugar:	$y = -0.0001 x^2 + 0.12 x + 0.95$	$r^2 = 0.95$	(4)

All relations presented correlation coefficients higher than 0.93. The x notation indicates the concentration of contamination $(g.kg^{-1})$ (independent variable), while the y notation indicates the relative area occupied by the adulterant in the RGB image (dependent variable). For instance, Eq. (1) shows that a result of 50% obtained by the supervised classification corresponds to a concentration of contamination by coffee husks and straw of 175 g.kg⁻¹. A relative area greater than 5.45% in the same equation means that the sample will

be considered adulterated because it will contain a concentration of coffee husks and straw greater than 10 g.kg⁻¹, which is the maximum amount of impurity tolerated by the Brazilian government. In the same way, Eq. (2) shows that 20% corresponds to a concentration of contamination by maize of 329 g.kg⁻¹. Significance tests (t-test) obtained by the SAS statistical package (SAS 1988) showed that all relations were statistically significant at $\alpha = 0.01$.



FIG. 1. TYPICAL RED/GREEN/BLUE (RGB) COLOR COMPOSITE OF PURE (a) AND CONTAMINATED COFFEE WITH COFFEE HUSKS (b)



FIG. 2. SPECTRAL SEPARABILITY AMONG BROWN SUGAR (A), MAIZE (B), PURE COFFEE (C), SHADOW (D), SOYBEAN (E) AND STRAW (F) The maximum range of DN is from 0 to 255.



(a)



FIG. 3. RELATION BETWEEN CONCENTRATION OF COFFEE HUSKS & STRAW (a) AND MAIZE (b) AND THEIR RELATIVE AREA IN THE RED/GREEN/BLUE (RGB) COLOR COMPOSITE



(a)



(b)

FIG. 4. RELATION BETWEEN CONCENTRATION OF SOYBEAN (a) AND BROWN SUGAR (b) AND THEIR RELATIVE AREA IN THE RED/GREEN/BLUE (RGB) COLOR COMPOSITE

Regression Analysis

The dataset prepared by the ABIC to test the method presented the following types and amount of adulterants (Tables 1 and 2): coffee husks & straw (25-400 g.kg⁻¹); maize (25-450 g.kg⁻¹); brown sugar (50 and 100 g.kg⁻¹) + maize (100 and 200 g.kg⁻¹); and coffee husks & straw (10 and 50 g.kg⁻¹) +

soybean (50 and 100 g.kg⁻¹). The total number of samples analyzed for each class of contaminant were: 90 for coffee husks & straw; 90 for maize; 28 for brown sugar + maize; and 29 for coffee husks & straw + soybean. The

TABLE 1.
REGRESSION ANALYSIS OF LABORATORY-PREPARED SAMPLES WITH VARYING
AMOUNTS AND ONE TYPE OF COFFEE ADULTERANT PER SAMPLE

Contaminant	Number of	Amount (g.kg ⁻¹)		Correlation
	Samples	Actual	Measured	Coefficient
Coffee Husks	10	2.5	2.9	0.972
& Straw	10	5	5.5	
	10	10	10.4	
	10	15	14.9	
	10	20	19.6	
	10	25	26.2	
	10	30	29.2	
	10	35	34.2	
	10	40	38.8	
Maize	10	25	28.2	0.991
	10	50	53.3	
	10	100	103.0	
	10	150	153.4	
	10	200	201.3	
	10	250	252.4	Í
	10	300	299.6	
	10	400	408.6	
	10	450	472.9	

TABLE 2.
REGRESSION ANALYSIS OF LABORATORY-PREPARED SAMPLES WITH VARYING
AMOUNTS AND TWO TYPES OF COFFEE ADULTERANT PER SAMPLE

Contaminant	Number	Amount (g.kg ⁻¹)				Correlation
	of	Actual		Measured		Coefficient
	Samples	Impurity 1	Impurity 2	Impurity 1	Impurity 2	
Brown Sugar	7	50	100	50.5	98.5	0.997
+ Maize	7	50	200	50.3	203.6	
	7	100	100	103.1	107.8	
	7	100	200	101.4	199.3	
Coffee Husks	7	10	50	10.7	59.7	0.898
& Straw +	7	10	100	10.4	96.1	
Soybean	7	50	50	51.6	56.2	
	8	50	100	44.5	103.1	

correlation coefficients clearly show the valuable performance of the new method: 0.972 for coffee husks & straw; 0.991 for maize; 0.998 for brown sugar + maize; and 0.898 for coffee husks & straw + soybean. All relations were statistically significant at $\alpha = 0.01$. A set of 10 additional pure coffee samples were also included in the dataset, which were 100% correctly classified by the visual analysis in the stereo microscope.

CONCLUSIONS

The results reported herein indicate that a new method based on digital image processing has facilitated the quantifying of the amount of contaminants in roast and ground coffees. Using laboratory-prepared samples with varying amounts and types of coffee substitutes as test materials, correlation values ranging from 90% to 99% were achieved. This new method presents three major desired properties. It is neither time consuming nor destructive and it is able to quantify a wide range of adulterant concentrations. The high correlations obtained for samples with two different contaminants also indicate that the method can work well when more than one adulterant is present in the sample. In other words, this method is suitable for routine analyses and the transferring of this technique for commercial use can be immediate.

ACKNOWLEDGMENTS

The authors wish to acknowledge the partial financial support of the Brazilian Association of Coffee Industries (ABIC). We are also thankful to FUNCAFE for providing a scholarship to the third author.

REFERENCES

- BRASIL. 1999. Portaria no. 377, de 26/04/1999. Regulamento ténico para fixação de identificação e qualidade de café torrado em grão e café torrado e moído. Diário Oficial da União [da República Federativa do Brasil] (http://www.anvisa.gov.br/legis/portarias/377 99.htm) (In Portuguese).
- BRIANDET, R., KEMSLEY, E.K. and WILSON, R.H. 1996. Discrimination of Arabica and Robusta in instant coffee by Fourier transform infrared spectroscopy and chemometrics. J. Agric. Food Chem. 44(1), 170-174.
- CAMARA, G., FREITAS, U.M., SOUZA, R.C.M. and GARRIDO, J. 1996. SPRING: Integrating remote sensing and GIS by object-oriented data modeling. Computers Graphics 15(6), 13-22.
- CAMPBELL, J.B. 1987. Introduction to Remote Sensing. 551 p. The Guilford Press, New York.
- INSTITUTO ADOLFO LUTZ. 1985. Normas analíticas do Instituto Adolfo Lutz. Determinação de umidade por Karl Fischer, Extrato Aquoso e Nitrogênio Total, 1, 3^a Ed. (In Portuguese), Sao Paulo.
- KNIGHT, P. 1997. Brazilian coffee scene. Planters' Chronicle 92(6), 249-257.
- MADEIRA NETTO, J.S. 1996. Spectral reflectance properties of soils. Photo-Interpretation 1996/2, 59-70.
- MENEZES JR., J.B.F. and BICUDO, B.A.A. 1958. On a microscopic method to quantify coffee husks in powdered coffee. Rev. Inst. Adolfo Lutz 11, 13-47 (In Portuguese).
- SABINS, F.F. 1997. Remote Sensing: Principles and Interpretation. 3rd Ed., Chap. 11, pp. 361-385, W.H. Freeman and Co., New York.
- SAS Institute. 1988. SAS/STAT User's Guide, Re/ease 6.03 Ed., pp. 1028, The SAS Institute, Cary, NC.