Classification of various grains using optical properties

S. MAJUMDAR, D.S. JAYAS, J.L. HEHN and N.R. BULLEY

Department of Biosystems Engineering, 438 Engineering Building, University of Manitoba, Winnipeg, MB, Canada R3T 5V6. Received 24 August 1995; accepted 11 March 1996.

Majumdar, S., Jayas, D.S., Hehn, J.L. and Bulley, N.R. 1996. Classification of various grains using optical properties. Can. Agric. Eng. 38:139-144. Rapid identification of the principal constituents of bulk grain samples can be used to automate the movement of grain through a terminal or processing plant and can also be important for quality control. This paper examines the potential of using tristimulus values, color solid scale values, and reflectance characteristics of grains as discriminating features among various seed types from bulk samples of cereals, pulses, oilseeds, and speciality crops. Visible light between 400 and 700 nm was used to collect data for separation of various grains. Most of the grains could be distinguished from each other using X, Y, and Z tristimulus values; L, a, and b color solid scale values; or percent reflectance ratios and ratios of slopes of the percent reflectance curves at different wavelengths. The bulk reflectance measurement provides a rapid first stage identification and narrows the possibilities for a second, more detailed, identification stage. Keywords: color, spectrocolorimeter, identification, reflectance, grain grading, tristimulus values.

L'identification rapide des principaux constituants d'un échantillon de grains peut être utilisée pour automatiser les opérations d'un terminal ou d'une usine de transformation et peut s'avérer importante pour le contrôle de la qualité. Cet article examine la possibilité d'utiliser les composantes trichromatiques, une échelle de couleur continue et les caractéristiques de réflectance des grains afin de différencier les types de grains d'échantillons de céréales, de légumineuses à grain, d'oléagineux et autres produits particuliers. On a utilisé la lumière visible entre 400 et 700 nm afin de recueillir des données sur la séparation de plusieurs types de grains. On a réussi à différencier la plupart des grains en utilisant des valeurs trichromatiques de X, Y et Z; des valeurs de L, a et b pour l'échelle de couleur: ou des valeurs de réflectance en pourcentage et des pentes de courbes de réflectance pour différentes longueurs d'onde. La mesure de la réflectance d'échantillons de grains est une méthode rapide d'identification et permet de réduire les possibilités lors d'une deuxième étape d'identification plus détaillée. Mots-clés: couleur, spectrocolorimètre, réflectance, identification, classement des grains, composantes trichromatiques.

INTRODUCTION

The grain industry has long relied on human judgement to analyze grain samples for those particles of foreign materials which are not machine-separable. Grain grading is primarily based on a subjective interpretation of grading factors that influence end-use quality. Grain inspectors determine the grade by visual inspection and assessment of grain samples relative to established standard grain samples. The lack of an objective method of measuring the important grading factors (e.g. varietal purity, vitreousness, and soundness) has led to inconsistent grade determinations to the extent that in Canada 10-14% of reinspected samples receive a different grade from

that originally assigned (Bevilacqua 1987). Therefore, a device that would use some physical properties of grains to separate kernels of a particular grain from other grains and foreign materials which are at present not machine-separable would be desirable. Such equipment would be extremely valuable to the grain industry for use in quality control. Digital image analysis can be effectively used for objective measurement of some of the grain quality parameters. The primary reason for its potential application for grain inspection and grading lies in its ability to quantify with precision, speed, and consistency the composition and physical characteristics of grain samples using parameters which form the basis of visual inspection (e.g. object size, shape, color, reflectance, and texture) (Sapirstein and Bushuk 1989). In this paper we discuss the feasibility of using the optical properties of grains as the basis for separating various types of grains.

Massie and Norris (1965) evaluated the reflectance and transmittance properties of grains for designing infrared grain dryers. They found that the moisture content did not have any significant effect on reflectance and transmittance properties of different grains and reflectance was the most important variable for infrared drying.

Hawk et al. (1970) measured the reflectance characteristics of various grains using a spectroreflectometer in ultraviolet, visible, and infrared regions of the spectrum to develop a model for grain separation. They found that the greatest difference in reflectance among various grains occurred between 450 and 700 nm. Also, they could not distinguish hard red spring (HRS) wheat from hard red winter wheat or barley from oats on the basis of reflectance data.

Dowell (1992) investigated the potential of tristimulus values and spectral reflectance for identification of normal and damaged peanut kernels using a monochromatic machine vision system, a contact colorimeter, a non-contact colorimeter, and a spectrophotometer and achieved 62.7, 79.1, 85.8, and 94.1% correct classification, respectively.

Casady et al. (1993) measured the optical properties of soybean seeds using a spectroradiometer. They found that the reflectance between 436 and 724 nm provided the most information for separation of normal from damaged soybean seeds.

OBJECTIVES

The present work is a part of a project to design and fabricate a machine vision system to grade various grains at grain handling facilities. As part of the development of image processing techniques using reflectance and color features, optical properties of various grains need to be studied. The specific objectives were:

- to test the feasibility of using the tristimulus values and the color solid scale values of bulk samples of various seed types for separation of one species of grain from other grains, and
- 2) to determine the percent reflectance of bulk samples of various seed types and to test the feasibility of using percent reflectance ratios and ratios of slopes of the percent reflectance curves at different wavelengths for their identification.

MATERIALS AND METHODS

Instrumentation

A spectrocolorimeter (LabScan 6000, Hunter Associates Laboratories, Inc., Reston, VA) was used to determine the percent reflectance at wavelengths between 400 and 700 nm with 10 nm increments for bulk samples of various seed types. The instrument also gave an indication of color in the form of XYZ tristimulus values (1964 CIE tristimulus values of a color in D₆₅ on the basis of Y = 100 for white) and Lab color solid scale values.

The value of X represents the amount of a reddish primary that has higher saturation than any obtainable red. The value of Y represents the amount of a green primary that is considerably more saturated but of the same hue as the spectrum color whose wavelength is 520 nm. The value of Z represents the amount of a blue primary that is considerably more saturated but of the same hue as the spectrum color whose wavelength is 477 nm. The tristimulus values can be translated, as given below, into three-dimensional sets of color space coordinates which indicate color as perceived by the human eye (MacAdam 1985):

$$R = 0.7990 X + 0.4194 Y - 0.1648 Z \tag{1}$$

$$G = -0.4493 X + 1.3265 Y + 0.0927 Z$$
 (2)

$$B = -0.1149 X + 0.3394 Y + 0.7170 Z$$
 (3)

Red (R), green (G), and blue (B) describe three dimensions of color. These dimensions can also be expressed in the form of intensity (I), saturation (S), and hue (H) (Ballard and Brown 1982) as:

$$I = R + G + B \tag{4}$$

$$S = 1 - \frac{3 \min(R, G, B)}{I}$$
 (5)

$$H = \cos^{-1} \left\{ \frac{\frac{1/2[(R-G) + (R-B)]}{\sqrt{(R-G)^2 + (R-B)(G-B)^{1/2}}}}{\sqrt{(R-G)^2 + (R-B)(G-B)^{1/2}}} \right\}$$
 (6)

Intensity indicates how bright the color is, such as light green. Saturation indicates how much of the color is there, i.e. purity of color. Hue indicates what color, such as green, dominates the reflected light.

The L value is an indication of brightness and can be compared to the reflectance information. The amount of red

and green is indicated by a positive or negative a value, respectively. Similarly, the b value indicates the yellow and blue color in the sample with its respective positive or negative number (Mohsenin 1984). Francis (1987) suggested that the a and b values could be reduced to a single measure θ , the angle between a line joining the point (a, b) in the Hunter space with the green-red axis as:

$$\theta = \tan^{-1} \left(b/a \right) \tag{7}$$

The value θ then becomes an independent variable and provides an indication of hue.

The spectrocolorimeter is comprised of an optical sensor with a sensor interface unit and an IBM personal computer. The optical sensor used 0° incident light on the sample plane. Viewing was at 45° through a ring of 16 fibre optic receptor stations. The light picked up by the fibre optic receptor module passed through a circular variable filter (CVF) scanner which was driven by a motor. The CVF, a monochromator, was used to separate light into its spectral components. The CIE illuminant D_{65} and 10° standard observer were used to calibrate the meter. For standardization of the meter, reflectance from a black glass was used as zero and reflectance from a white calibrated standard was set at 100%.

Sample preparation and measurement procedure

The spectrocolorimeter was allowed to warm up for a period of 1 h for temperature equilibration. For the reflectance measurements, a clear plastic petridish (63.5 mm dia x 12.7 mm deep) filled with the sample was placed on the reflectance port of the spectrocolorimeter. Table I shows various types of grains that were tested using the instrument. Five samples from each grain type were tested. All grains were conditioned to a specific moisture content (13.5±0.5% wet basis).

The 'Random Cup Method' sampling procedure was used for sample preparation (Anonymous 1992). Five petridishes were placed on a tray. After a preliminary pan-to-pan mixing and then passing the grain through a *Boerner Divider* for 3-4 times for thorough mixing, the grain was poured uniformly over the tray. The excess material was removed by once passing a straight edge over the top of the samples. The grains that fell into the petridishes were taken as the working samples. Three images of each sample were taken. After acquiring one image the sample was re-mixed and the next image was taken. Similarly, the third image was taken.

RESULTS AND DISCUSSIONS

Feature set

The percent reflectance, over the visible spectrum, for various grains is shown in Figs. 1 to 4. None of the Canadian western red spring (CWRS) wheat grades could be distinguished solely based on their reflectance characteristics (Fig. 1). It would be very difficult to utilize the information gathered from the percent reflectance data for practical purposes because the percent reflectance of any object can change due to changes in many parameters like intensity of light, image background, dust, aging of the light sources, etc. This problem could be minimized to a great extent by using the ratios of slopes of percent reflectance curves and the ratios of percent reflectance at different wavelengths:

Table I: Tristimulus values of various grains obtained from LabScan 6000 Spectrocolorimeter*

Seed type	X		Y		Z	
	Mean	S.D.	Mean	S.D	Mean	S.D.
		Cereals				
HRS wheat, grade 1	20.47 EFG	0.49	19.73 FGH	0.55	10.58 FGHI	0.49
HRS wheat, grade 2	20.13 EFG	0.42	19.50 FGH	0.44	10.72 FGHI	0.38
HRS wheat, grade 3	20.33 EFG	0.33	19.71 FGH	0.32	10.92 FGH	0.16
HRS wheat unclean grade 1	20.19 EF	0.76	20.38 EFGH	0.83	11.62 FGH	0.64
IRS wheat unclean grade 2	19.42 EFG	0.34	18.69 GH	0.36	10.13 GHI	0.36
IRS wheat unclean grade 3	19.30 FG	0.51	18.85 GH	0.45	10.77 FGHI	0.17
Durum wheat	21.83 DE	0.12	21.30 DEFG	0.69	10.55 FGHI	0.50
Canada prairie spring (CPS) wheat	23.99 CD	0.90	23.46 CD	0.93	11.80 EFG	0.49
frow (6R) barley (Malt)	32.36 AB	0.63	32.92 AB	0.66	18.18 B	0.34
Feed barley	25.34 C	0.49	25.54 C	0.46	13.79 DE	0.21
Dats	30.28 B	0.45	30.72 B	0.28	17.09 BC	0.14
Rye	18.23 G	0.23	18.29 H	0.27	11.05 FGH	0.25
		Pulses				
White pea bean	33.98 A	1.19	35.11 A	1.26	27.79 A	1.24
Pinto bean	20.44 EFG	1.19	19.79 FGH	1.23	13.69 DE	1.08
Black bean	3.22 K	0.17	3.36 M	0.16	3.73 K	0.16
Field pea-green seeded	19.92 EFG	0.93	21.50 DEF	0.86	15.06 CD	0.96
Dark green speckled' (DGS) lentils	7.84 I	0.48	8.17 JF	0.48	6.26 J	0.20
Eston' lentils	18.49 FG	0.12	18.34 H	0.20	9.71 HI	0.18
Laird' lentils	20.70 EF	0.89	20.89 EFG	0.79	12.40 EF	0.41
	Oilsee	ds and specialit	y crops			
Yellow mustard	23.74 CD	0.24	22.49 DE	0.24	10.14 GHI	0.18
Driental mustard	20.52 EFG	0.60	19.40 FGH	0.58	6.35 J	0.18
Brown mustard	5.01 JK	0.13	4.67 LM	0.11	3.77 K	0.01
Sunflower seed	10.67 H	0.98	11.13 I	0.93	8.82 I	0.96
Canola	4.17 JK	0.09	3.86 M	0.08	2.88 K	0.09
Flax seed	10.53 H	0.22	9.86 IJ	0.22	6.02 J	0.13
Buckwheat	6.66 IJ	0.61	6.56 KL	0.64	4.72 J	0.43

^{*} Means in columns followed by the same letter(s) are not significantly different at P = 0.05.

Slope – ratio 1 =
$$\frac{\% \text{ reflectance at 670 nm} - \% \text{ reflectanceat 620 nm}}{\% \text{ reflectance at 570 nm} - \% \text{ reflectance at 520 nm}}$$
(8)

% reflectance at 500 nm – % reflectance at 450 nm

$$Slope - ratio 3 =$$

$$Ratio 1 = \frac{\% \ reflectance \ at 650 \ nm}{\% \ reflectance \ at 550 \ nm}$$
 (11)

$$Ratio 2 = \frac{\% \ reflectance \ at \ 600 \ nm}{\% \ reflectance \ at \ 500 \ nm}$$
 (12)

$$Ratio 3 = \frac{\% \ reflectance \ at 550 \ nm}{\% \ reflectance \ at 450 \ nm}$$
 (13)

The selection of the wavelengths in Eqs. 8-13 was done intuitively by observing the trends of graphs in Figs. 1 to 4. The tristimulus values (X, Y, and Z), color solid scale values (L, a, and b), ratios of slopes of the percent reflectance curves (Slope-ratio 1, Slope-ratio 2, and Slope-ratio 3), and percent reflectance ratios (Ratio 1, Ratio 2, and Ratio 3) were tested for their potential in identification of various grains. Scheffe's multiple comparison test (SAS 1990) was carried out using the tristimulus values (Table I) and color solid scale values (Table II) of various seed types at 95% confidence level.

Tristimulus values

The mean X tristimulus values of the three grades of the CWRS wheat were significantly different from other cereal

(9)

grains except that of the durum wheat because the reddish primary of CWRS wheat and durum wheat overlap each other. The mean X, Y, and Z tristimulus values of the six-row barley were different from all other cereal grains except those of oats grains because the brightness of six-row barley and oats was very similar. The white pea bean could be separated from other grains using any tristimulus value because it was the brightest of all grains. The mean X, Y, and Z tristimulus values of the black bean and the 'Dark green speckled' (DGS) lentils were different from other pulse crops. The mean Z tristimulus value of the 'Eston' lentils was different from other pulse crops.

The mean X tristimulus value of the yellow mustard was different from other oilseeds and speciality crops as its color was different. The X tristimulus values of the sunflower seed and the flax seed were different from other oilseeds and speciality crops but they were not different from each other.

Solid scale values

The L color scale values of the six-row barley and the oats were different from other cereal grains but they were not different from each other as their brightness is very similar to each other (Table II). The a color scale values of all wheat varieties were also significantly different from other cereal grains.

Using the L color scale value, the white pea bean, the black bean, and the 'DGS' lentils could be distinguished from each other as well as from other pulse crops. The a color scale values of the white pea bean, the pinto bean, and the field pea were different from each other as well as from other pulse crops.

The brown mustard and the canola could not be separated from each other using L, a, and b values. Using b color scale values, all the oilseeds and speciality crops could be differentiated from each other except the brown mustard and the canola, and sunflower seed and buckwheat.

Table II: Color solid scale values of various grains obtained from LabScan 6000 spectrocolorimeter*

Seed type	$oldsymbol{L}$		а		b	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
		Cereals				
HRS wheat, grade 1	44.41 FGH	0.62	7.20 CD	0.22	14.84 E	0.11
HRS wheat, grade 2	44.15 FGH	0.50	6.77 CD	0.20	14.37 EF	0.24
IRS wheat, grade 3	44.39 FGH	0.36	6.68 CD	0.13	14.33 EF	0.15
IRS wheat unclean grade 1	45.14 EFGH	0.91	6.38 CD	0.36	14.13 EF	0.24
IRS wheat unclean grade 2	43.23 GH	0.41	7.11 CD	0.13	14.28 EF	0.11
IRS wheat unclean grade 3	43.40 FGH	0.52	5.97 DE	0.35	13.53 FG	0.46
Ourum wheat	45.82 DEGH	0.15	7.60 BC	0.08	16.55 D	0.13
Canada prairie spring (CPS) wheat	48.43 CD	0.96	6.54 CD	0.13	17.14 CD	0.44
row (6R) barley (Malt)	57.37 AB	0.57	3.63 FGH	0.07	18.59 B	0.26
eed barley	50.54 C	0.46	4.01 FGH	0.24	16.76 CD	0.40
Dats	55.42 B	0.26	3.79 FGH	0.66	17.81 BC	0.34
Rye	42.76 H	0.32	3.35 H	0.32	12.46 G	0.26
		Pulses				
White pea bean	59.25 A	1.06	2.08 I	0.09	10.40 H	0.16
Pinto bean	44.47 EFGH	1.38	6.84 CD	0.39	10.56 H	0.44
Black bean	18.34 M	0.44	0.25 K	0.22	-0.42 L	0.09
Field pea-green seeded	46.36 DEF	0.94	-1.83 L	0.59	10.76 H	0.29
Dark green speckled' (DGS) lentils	28.83 J	0.72	0.64 K	0.11	5.59 J	0.49
Eston' lentils	42.82 H	0.24	4.69 FG	0.59	14.46 EF	0.15
Laird' lentils	45.69 DEFGH	0.86	3.53 GH	0.50	13.63 EFG	0.37
	Oilseeds	and specialit	y crops			
Yellow mustard	47.43 DE	0.25	9.24 A	0.16	18.35 B	0.09
Oriental mustard	44.05 FGH	0.66	8.71 AB	0.08	20.43 A	0.35
Brown mustard	21.62 L	0.24	4.86 EF	0.26	3.56 K	0.27
Sunflower seed	33.09 I	1.57	1.47 JK	0.24	5.56 J	0.33
Canola	19.65 LM	0.21	4.68 FG	0.19	3.99 K	0.26
Flax seed	31.40 IJ	0.34	6.81 CD	0.05	9.04 I	0.11
Buckwheat	25.58 K	1.21	3.16 HI	0.23	5.61 J	0.47

^{*} Means in columns followed by the same letter(s) are not significantly different at P = 0.05.

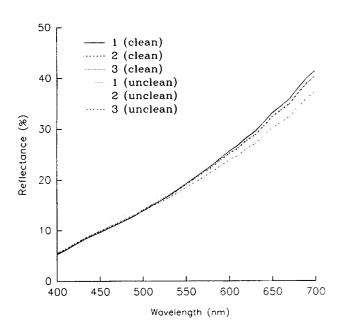


Fig. 1. Reflectance characteristics of different grades of hard red spring wheat (HRSW).

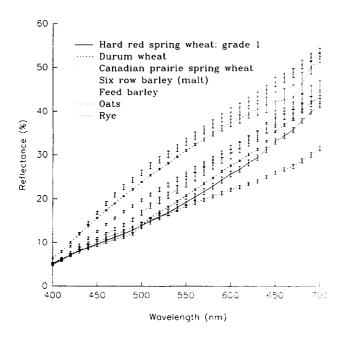


Fig. 2. Confidence interval comparisons (95%) of reflectance of various cereal grains.

Discriminant analysis

Although some grains could be separated using various features $(X \ Y \ Z \ and \ L \ a \ b)$, no single parameter could separate all different grains. To assess the pattern classification capacity of these features, procedure DISCRIM of SAS (1990) was used to classify various bulk grain samples. The $X \ Y \ Z$ tristimulus values, the $L \ a \ b$ color solid scale values, and the

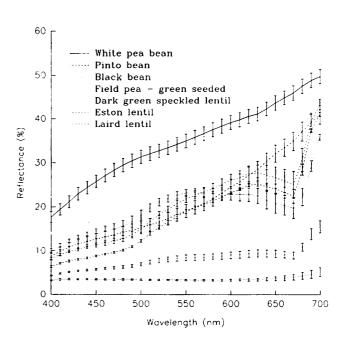


Fig. 3. Confidence interval comparisons (95%) of reflectance of various pulse crops.

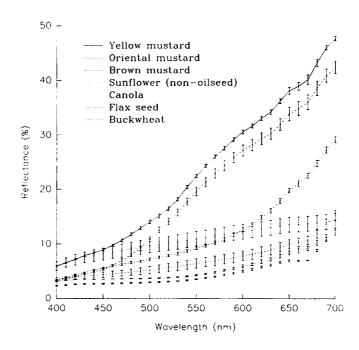


Fig. 4. Confidence interval comparisons (95%) of reflectance of various oilseeds and speciality crops.

Slope-ratios 1 to 3 and Ratios 1 to 3 features were separately used for discriminant analyses for classification of various grains. A non-parametric probability density estimation (e.g. k-nearest neighbour) was used with k value of 2. The squared distance between two observation vectors, x and y, in group t is given by:

 $d_t^{2}(x, y) = (x - y)' V_t^{-1}(x - y)$ (14)

where:

 $V_t = S$ the pooled covariance matrix,

x, y = p-dimensional vectors containing the quantitative variables of an observation, and

t = a subscript to distinguish the groups.

The classification was based on the Bayes decision rule which classifies an entity (represented by its pattern vector, e.g. x) to a class for which the entity has a maximum posteriori probability (Hand 1981; Duda and Hart 1973). An observation x is classified into group u if setting t=u produces the largest value of p(t/x). If there is a tie for the largest probability or this largest probability is less than the threshold specified, x is classified into group 'other'. In the leave-one-out classification method, one sample is considered as a testing set at a time and the remaining n-1 samples are used as a training set. This method yields the most unbiased and conservative estimate of the true error among all commonly used methods of error estimation (Hand 1981).

In the discriminant analyses, only grade 1 samples of the CWRS wheat were used. When the XYZ tristimulus values were used for analysis (leave-one-out method with 21-way classification), all samples of various grains were correctly classified except one sample of the sunflower seeds and the Canada prairie spring (CPS) wheat, which were misclassified as the 'DGS' lentils and the durum wheat, respectively.

CONCLUSIONS

The reflectance characteristics and tristimulus values of various grains were measured over the visible range (400 - 700 nm) of the electromagnetic spectrum. The tristimulus values, the color scale values, and the percent reflectance ratios and the ratios of slopes of the percent reflectance curves at different wavelengths could be effectively used to identify most of the grains used in this study.

PRACTICAL IMPLICATION

As there is a linear relationship between tristimulus values and R, G, and B values (Eqs. 1-3), a color image (RGB) can be used to calculate the tristimulus values that can be used for discriminant analysis for classification of different types of grains. Also, filters with specific wavelengths can be used in a camera to extract reflectance features at those wavelengths for discriminant analysis of different types of grains.

ACKNOWLEDGEMENTS

We thank Dr. S. Sokhansanj, Professor, Department of Agricultural and Bioresource Engineering, University of Saskatchewan for allowing us to use the spectrocolorimeter (LabScan 6000) and Prince Rupert Grain Ltd., the Natural Sciences and Engineering Research Council of Canada, Agriculture and Agri-Food Canada, and University of Manitoba Graduate Fellowship Committee for partial funding of this study.

REFERENCES

- Anonymous. 1992. Canadian methods and procedures for testing seed. Laboratory Services Division, Food Production and Inspection Branch, Agriculture Canada, Ottawa. ON.
- Ballard, D.H. and C.M. Brown. 1982. Computer Vision. Englewood Cliffs, NJ: Prentice-Hall, Inc.
- Bevilacqua, R.T. 1987. Grading of grains and oilseeds in Canada. Lecture presented at the 22nd International Grain Industry Course. Canadian International Grain Institute, Winnipeg, MB.
- Casady, W.W., M.R. Paulsen and J.B. Sinclair. 1993. Optical properties of damaged soybean seeds. *Transactions of the ASAE* 36(3):943-947.
- Dowell, F.E. 1992. Identifying undamaged and damaged peanut kernels using tristimulus values and spectral reflectance. *Transactions of the ASAE* 35(3):931-937.
- Duda, R.O. and P.E. Hart. 1973. Pattern Classification and Scene Analysis. New York, NY: John Wiley and Sons, Inc.
- Francis, F.J. 1987. Food Colorimetry: Measurement and Interpretation. Physical Properties of Foods-2. New York, NY: Elsevier Applied Science.
- Hand, D.J. 1981. Discrimination and Classification. New York, NY: John Wiley and Sons, Inc.
- Hawk, A.L., H.H. Kaufmann, and C.A. Watson. 1970. Reflectance characteristics of various grains. *Cereal Science Today* 15(11): 381-384.
- MacAdam, D.L. 1985. Color Measurement: Theme and Variations, 2nd ed. New York, NY: Springer-Verlag.
- Massie, D.R. and K.H. Norris. 1965. Spectral reflectance and transmittance properties of grain in the visible and near infrared. *Transactions of the ASAE* 8(4):598-600.
- Mohsenin, N.N. 1984. Electromagnetic Radiation Properties of Foods and Agricultural Products. New York, NY: Gordon and Breach Science Publishers.
- Sapirstein, H.D. and W. Bushuk. 1989. Quantitative determination of foreign material and vitreosity in wheat by digital image analysis. In *ICC '89 Symposium: Wheat End-Use Properties*, ed. H. Salovaara, 453-474. Lahti, Finland.
- SAS. 1990. SAS User's Guide: Statistics. Raleigh, NC: Statistical Analysis System, Inc.