Application of Pattern Recognition Techniques in the Analysis of Cereal Grains[1]

F. S. LAI,[2] I. ZAYAS,[3] and Y. POMERANZ[2]

Cereal Chem. 63(2):168–172

ABSTRACT

Objective methods are required to characterize grains in plant breeding programs and in marketing channels. Pattern recognition techniques can be used as an aid in such characterization. Pattern recognition, a learning algorithm stored in a computer, was found effective in identifying and classifying cereal grains. Patterns of six grains (corn, soybeans, rice, sorghum grain, barley, and wheat) were developed, and the patterns were tested for their accuracy in the recognition of grains. In addition, the technique was applied to differentiate between brown and white rice and in the study of sphericity of corn kernels.

The classification process of a pattern recognition system partitions an n-dimensional feature space into disjointed regions, each associated with only one class (Andrews 1972, Ahmed and Rao 1975, Chen 1973). The separating surfaces are often referred to as hyperplanes, or discriminant functions, and are of $n - 1$ dimensions. Thus, the hyperplane separating data on a plane of $n-2$ dimensions is a line. The easiest partition is linear and requires n computational steps for the classification process. However, it is often necessary to use the concept of nonlinear separating surfaces, because data from physical phenomena may not be linearly separable. Classification algorithms usually are designed to use some optimal criteria (Young and Calvert 1974, Jurs and Isenhour 1975, Martin and Reddy 1975). The partitions may be defined by deterministic and stochastic criteria, and the result is a segmentation of the n-dimensional feature space. This segmentation-simplified classification of a feature vector corresponds to an unknown pattern vector by observing its location in the partitioned space.

Pattern recognition techniques can be put to many potential uses. In plant breeding, the technique can be used to distinguish among classes and varieties and to determine the uniformity of kernel size and shape for a given cultivar. In grain inspection and quality control, the technique can be used to objectively analyze conformity with grain standards. In the processing industries, it can be used to measure foreign material, to determine uniformity of products, and to evaluate commercial value. In all cases, the technique provides an objective evaluation and a permanent record.

The U.S. Grain Marketing Research Laboratory, Agricultural Research Service, U.S. Department of Agriculture, Manhattan, KS, has a comprehensive program on the application of pattern recognition techniques in the analysis of cereal grains. We have encountered several problems in developing pattern recognition models. We were aware that it would be difficult to find the parameters that were broad enough to include all grains in the recognition of different grain types and at the same time specific enough to be meaningful for a single grain type. In addition, at this time, the cost of the equipment and the time required for analyses are prohibitive.

Those constraints notwithstanding, pattern recognition is an objective, nondestructive method that can provide a permanent record. Recent developments in instrumentation provide promise for automated, rapid, and less expensive equipment.

The objective of this paper is to provide a description of the approaches we used to identify different types of grain. Those approaches were then used to differentiate between brown and white rice and among corn kernels that varied in size and shape. The use of pattern recognition techniques to identify wheat classes and varieties will be described elsewhere (Zayas 1985).

MATERIALS AND METHODS

Materials

Grain kernels were picked randomly from samples from which broken, immature, and shrunken kernels were first removed by hand. The grains (hard red winter wheat, six-rowed barley, dent corn, grain sorghum, soybeans, and brown and white rice) were from commercial channels and were provided by the Federal Grain Inspection Service of the U. S. Department of Agriculture. The seven fractions of dent corn (hybrid Bojac) were separated mechanically by size and shape and were described elsewhere (Pomeranz et al 1985).

Methods

An optical framework, composed of a specific number of kernels, was made for each type of grain. The number of kernels that constituted a set depended upon the number of kernels that fit the screen size. For corn, five sets of 15 (75 kernels) were used to establish an image pattern for corn. The pattern was then evaluated against another set of 450 kernels of the hybrid Bojac. For wheat and barley, three sets of 36 kernels each were made; three sets of 20 kernels each were made for soybeans; two sets of 34 kernels each were made for rice; and three sets of 30 kernels each were made for sorghum grain.

All sets of grain kernels were kept in the same orientation; for example, corn kernels were placed lengthwise along the y-axis (parallel to the vertical side of the screen) with the germ up. For soybeans, black spots were placed downward.

A Quantimet 720 image analyzer (Cambridge Imanco, Monsey, NY) transferred data to a DEC PDPII computer via a DMA (direct memory access) interface, which also transferred software-generated displays back to the analyzer. The entire unit was linked to an IBM 370 via a CMS interface.

The operating system had access to all system facilities in an immediate mode, when results of the proposed measurement were seen at once, or in a program mode, when instructions were entered into a measurement routine. The image analyzer projected the image of a grain kernel through a camera (vidicon) and the image was digitized. The image was detected by a gray-level threshold, and an electronic bright-up display was superimposed over

[1]Presented at the AACC 69th Annual Meeting, Minneapolis, MN, September 1984. Cooperative investigations, U. S. Department of Agriculture, Agricultural Research Service, U.S. Grain Marketing Research Laboratory, Manhattan, KS 66502 and the Department of Chemical Engineering, Kansas State University. Contribution 85-340-J, Department of Chemical Engineering, Kansas Agricultural Experiment Station, Manhattan, KS 66506.

Mention of firm names or trade products does not constitute endorsement by the U. S. Department of Agriculture over others not mentioned.

[2]Research chemical engineer and research chemist, respectively.

[3]Research assistant, Department of Chemical Engineering, Kansas State University.

features, which were defined as any object detected by the image analyzer. Feature values were then determined and stored for subsequent calculation.

We developed several methodological strategies in the preparation of samples and in the measurement of features. The quality of the image of kernels selected at random was carefully controlled by selecting the gray level detection and by manipulating the illumination. The illumination provided an even light without shadow around objects on a nonreflective background. Scanner light sensitivity was controlled by peak white level, aperture opening, and illumination. To increase the image resolution, we used a lens with a focal length of 75 mm and 4× object magnification. To improve consistency and reliability of measurements, a calibration sphere close in color to the grain was used as a reference in each test.

Computer programs were developed to accept an image and calculate morphological parameters. Included in the programs were the capabilities to readjust the detection level in the programming mode and to check the labeling of successive members in each set of kernels. Checking was critical, because our method was based on the measurement of two orthogonal projections and the relationship between geometrical characteristics in two positions. Great care was exercised to ensure the correct orientation of each kernel towards the crease line and to keep sequential counts of the kernels according to their orientation toward the vertical axis.

The development of a simple recognition pattern was complicated by the unevenness of color of individual kernels, the subjective determination of kernel position and orientation, and the wide scattering of geometrical parameters for a specific type of grain. Variations in characteristics of biological material can be attributed to location, year, and variety, which reflect the complexity of the genotypic and phenotypic factors and their interactions.

Several parameters can be selected, singularly or in combination, to distinguish one shape from another. The basic parameters from the analyzer are shown in Figure 1 (Imanco 1978, Gonzalez and Wintz 1977, Loeb 1984, Attle et al 1980):

Area: Sum of all pixels within feature boundaries or detected phases. The pixels are the points, the basic picture elements. The output signal from vidicon after amplification is converted to a digital number corresponding to transmittance of a specific sample point or pixel in the image. In this study, the maximum numbers of pixels in a frame was 600,000.

Perimeter: Sum of all pixels on the boundary of a selected feature.

Projected height: Sum of scan line intercepts on a feature's horizontal trailing edge, also called 90° feret.

Projected width: Sum of scan line intercepts on a feature's vertical trailing edge, also called 0° feret.

Length: Maximum linear feature dimension independent of orientation.

Width: Feature dimension measured at 90° to longest dimension.

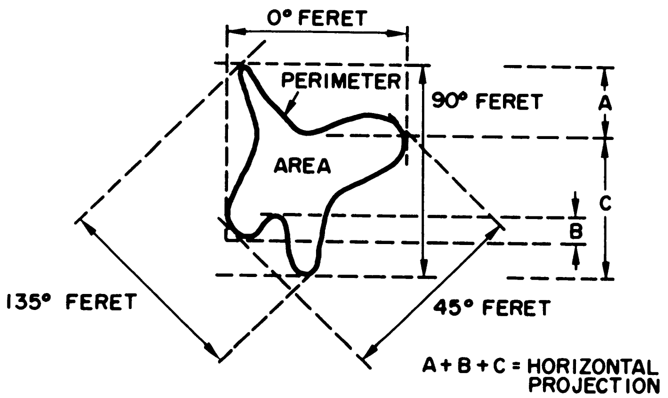

Fig. 1. Basic parameters of the image analyzer.

Feret: The substandard feature length in a defined vector input from the keyboard. The possible feret measurement in the image analyzer can be 0°, 45°, 90°, and 135°.

Volume: Integrated optical brightness or density within the feature boundary. This measurement is the reflectance of color.

TABLE I
Derived Diameters Available From Function Analyzer Parameters

Function Analyzer Parameters	Derived Diameters	Proportional To:
Area, A	$D1 = 2\sqrt{(A/\pi)}$	Area
Perimeter, Pe	$D2 = Pe/\pi$	Perimeter
Feret diameters, measured in n directions, F1, F2, ..., Fn	Mean feret (arithmetic) $D3 = \sum_{i=1}^{n} Fi/n$	Convex perimeter
Feret diameters, measured in n directions, F1, F2, ..., Fn	Mean feret (geometric) $D4 = n\sqrt{(F1 \cdot F2 \cdot F3 ... Fn)}$	Envelope area
Feret diameters, measured in n directions, F1, F2, ..., Fn	Maximum feret D5 = max (Fi)	Length
Feret diameters, measured in n directions, F1, F2, ..., Fn	Minimum feret D6 = min (Fi)	Envelope width

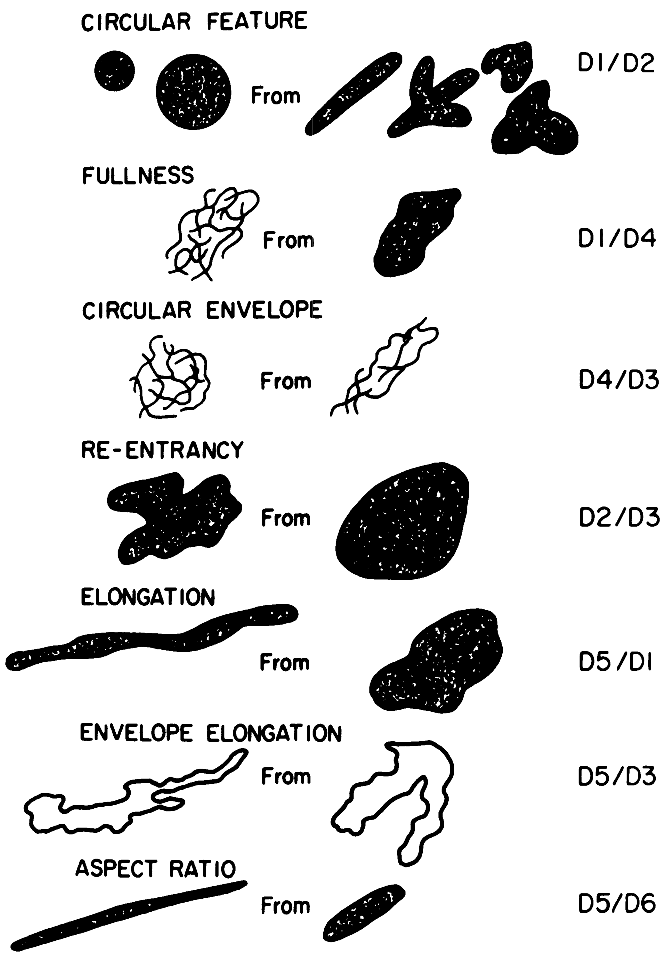

Fig. 2. Various shape functions.

Convex perimeter: Equivalent to the length of a string around the outside of a feature.

Count: number of features.

A detailed description of various parameters, based on the above measurements, is given in Table I (adapted from ASTM 1980 and Imanco Instruction Manual 1979). The first order shape function to distinguish among different shapes of particles is given in Figure 2. In all patterns, pixels were used as measuring units. Values of all criteria in patterns are not absolute and depend on the setting of the scene.

RESULTS AND DISCUSSION

The process of measuring the image of a particular type of grain is objective; however, the process of developing the pattern is subjective. The patterns we obtained were selected out of a great number of possible ones and were obtained by subjective judgment and by trial and error.

A major problem was the selection of variables to create a reliable pattern to characterize and distinguish a specific grain from others. After comparing simple geometrical values, such as area, perimeter, length, and width, we found that those measurements often were inadequate to identify a specific grain. Consequently, several shape functions were derived to establish the limits for each criterion for each type of grain.

Measurements of all basic parameters as well as derivative shape functions for all types of grains were graphed in histograms to establish ranges for these values. Minimum and maximum values for each specific parameter that was chosen were considered to be preset values. The preset values became the criteria to include parameters in the process of pattern recognition for each type of grain.

Corn

A set of kernels was measured in the germ-up position, which is the most stable position for a corn kernel. Of the six types of grain we studied, corn kernels were the largest. It was therefore logical to select the projected area as one of the pattern recognition criteria. The results after measuring the projected area of 75 kernels of commercial dent corn are given in a distribution histogram in Figure 3. The color of corn is also unique; consequently, measurement of the integrated optical brightness, called volume, was selected as another characteristic criterion for pattern recognition. The frequency histogram of volume (integrated optical brightness) is given in Fig. 4. Each parameter was individually compared with a preset value. With these two criteria, we developed one of several possible patterns for corn (Table II). This pattern was established after a test against six types of grain yielded a 100% accurate prediction.

Soybeans

Soybeans are distinguished from other grains by their relatively high degree of roundness, contributing to a difficulty in determining their plane of maximum stability. Both soybean and sorghum grains are round, but they differ widely in size, and there is no overlap, even for small soybean kernels. Consequently, we distinguished between the two by the simple criterion of projected area. However, because a small fraction of the larger soybean kernels closely resembled the smaller, rounder kernel fraction of corn, the dimensions of some soybean kernels overlapped with some spherical corn kernels. Therefore, we selected feret 0°/feret 135° as a measure of circularity and the projected area as a criterion. Based on these two criteria, we established the pattern for soybeans given in Table II.

Sorghum

Sorghum grains are similar to soybeans in shape, but differ

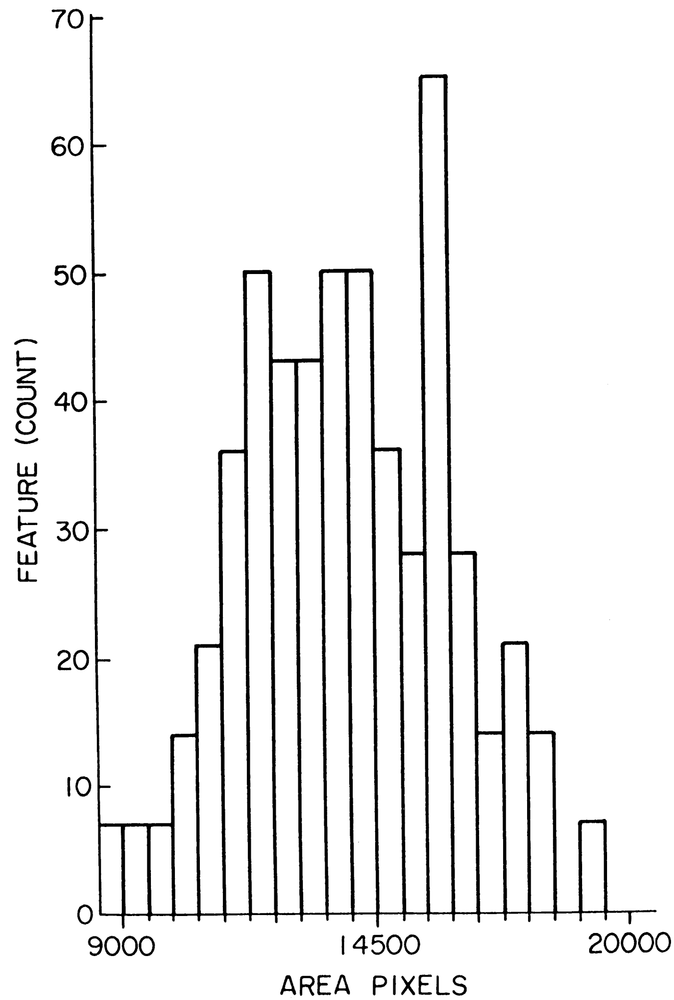

Fig. 3. Area histogram of corn kernels.

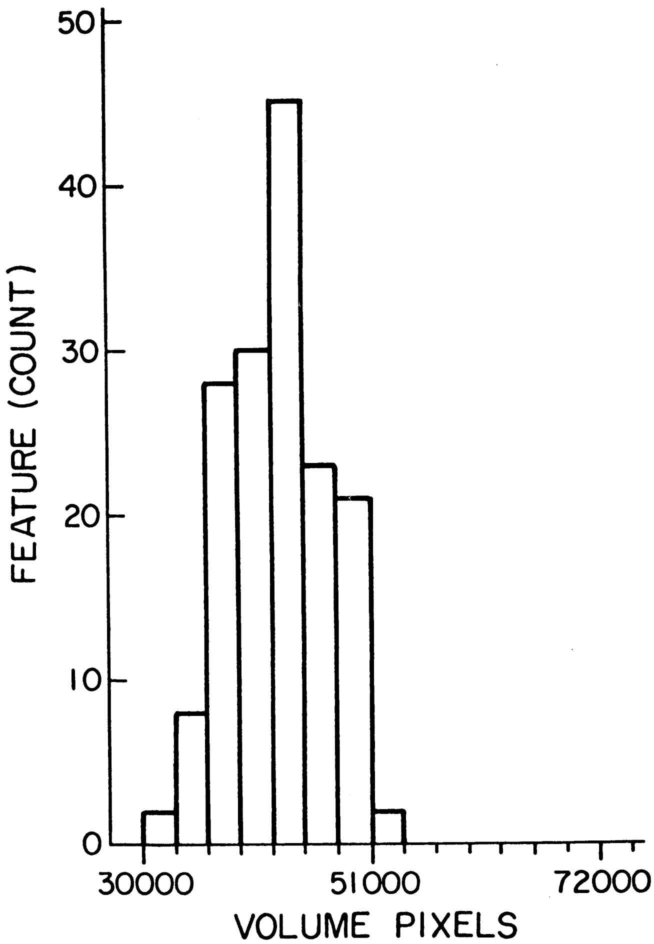

Fig. 4. Volume histogram of corn kernels.

significantly in size. To define the pattern for sorghum, we used the aspect ratio (length/width), area/volume, and feret 0°/feret 135°. The aspect ratio allowed us to separate a full shape from a thin and elongated shape. Using this parameter we separated sorghum from rice, wheat, and barley. Area/volume reflected the characterictics of projected area and color, which separated sorghum from soybeans and corn. The pattern is given in Table II.

Other Grains

We used the above approach and logic to develop patterns for the other grains (Table II). The convex perimeter and area/volume were the best criteria for brown rice. We used three criteria to establish the pattern for wheat. The pattern for barley was the most difficult to establish; nine criteria were needed for a reliable recognition.

As stated previously, the established criteria provided 100% reliable recognition for the samples used in this study. Additional criteria may be needed to distinguish closely related patterns such as for naked barley and wheat. In addition, colors for the various grains range from white to dark brown and include such off-colors as green, as well as nonuniform (at times mottled) color distribution. All these criteria must be considered in final and definitive characterization.

Milled Rice

Elongated and ellipsoid shapes of milled rice presented difficulties in determining the plane of maximum stability. Because of the distinctive white color of the milled rice, we used the ratio of area/volume as a function of size and color, and the combination of feret diameters and length/width as a characteristic of shape independent of the position of the kernel. The patterns for white and brown rice are given in Table II.

Analysis of Corn Kernel Shape

We extended our study on the identification of different types of grain to the shape analysis of corn kernels. Pomeranz et al (1985) studied the effect of size and shape on corn breakage susceptibility. They found consistent differences in gross composition or hardness among the fractions separated according to size and shape (sphericity). The corn samples were fractionated by a mechanical separator. In that study, length, thickness, and width were measured by calipers to determine sphericity. Sphericity is defined as:

$$\text{Sphericity} = (abc)^{1/3}/a,$$

where a = length or the longest intercept; b = width or longest intercept normal to a; and c = thickness, or longest intercept normal to a and b.

In this study, we placed the kernels under the image analyzer and recorded physical dimensions (Table III). The picture point was calibrated to critical dimensions, equivalent to 0.06666 mm. Corn kernels were measured twice, once with the germ facing upward (frontal projection), and once with the germ facing to the left (profile). The results for sphericity of the seven corn fractions are given in Table IV. We considered a coefficient of separation (mean value) to distinguish among the fractions. This coefficient is the ratio of the circularity shape factor (CSF) of frontal projection to that of the profile projection, i.e.,

$$\text{Coefficient of separation} = \text{CSF (frontal)}/\text{CSF (profile)},$$

where the CSF is defined as the median measurement of area/perimeter^2.

The correlation between circularity and sphericity was found to be 0.852. This value indicated that the correlation between the two parameters is significant at the 0.05 level. From Table IV, we can see that for individual kernels in flat and round subsets there is an overlap in coefficients of separation and sphericity; however, the mean values for flat and round did not overlap.

TABLE II
Pattern Parameters of Various Grains

Grain	Recognition Criteria Min		Max
Corn	9,500 <	Area	<20,000
	32,000 <	Volume	<60,000
Soybeans	7,000 <	Area	<9,950
	0.84<	Feret 0°/Feret 135°	< 1.1
Grain sorghum	0.8<	Length/Width	< 1.25
	0.12<	Area/Volume	< 0.20
	0.8<	Feret 0°/Feret 135°	< 1.2
White rice	50<	2 Area/π	< 66
	0.07<	Area/Volume	< 0.12
	0.3<	Feret 0°/Feret 135°	< 0.68
	2.2<	Length/Width	< 3.7
Brown rice	205<	Convex perimeter	< 260
	0.07<	Area/Volume	< 0.11
Wheat	0.45<	Feret 0°/Feret 135°	< 0 .74
	180<	Convex perimeter	< 245
	0.1<	Area/Volume	< 0.17
Barley	3,400<	Area	< 5,800
	0.08<	Area/Volume	< 0.18
	0.4<	Feret 0°/Feret 135°	< 0.68
	75<	Feret 135°	< 125
	1.9<	Length/Width	< 2.9
	26,000<	Volume	< 39,000
	2.5<	Perimeter/Width	< 7.2
	1.1<	Perimeter^2/(4 Area)	< 1.8
	180,000<	Length (Width/2)	< 370,000

TABLE III
Physical Dimensions of Seven Fractions of Bojac Corn

No.	Shape Description	No. of Kernels Tested	Projection Position	Mean Length (mm)	Mean Width (mm)	Mean Area (mm^2)
N1	Flat medium	80	Frontal	...[a]	7.36	65.32
		80	Profile	11.89	5.41	49.75
N2	Flat small	80	Frontal	12.52	7.51	72.25
		80	Profile	12.50	4.65	47.04
N3	Round large	70	Frontal	10.53	8.44	65.06
		70	Profile	...	7.45	56.72
N4	Flat round	70	Frontal	11.14	8.95	74.06
		70	Profile	11.15	6.43	53.33
N5	Flat large	70	Frontal	12.84	8.51	84.73
		70	Profile	12.77	4.89	49.78
N6	Round small	80	Frontal	11.13	7.42	62.91
		80	Profile	11.10	6.45	54.23
N7	Mixture	80	Frontal	10.60	6.27	50.61
		80	Profile	10.50	4.87	39.12

[a] Data not available.

TABLE IV
Circularity and Sphericity of Bojac Corn

No.	Shape Description	Coefficient of Separation Range	Coefficient of Separation Mean	Sphericity Range	Sphericity Mean
N1	Flat medium	0.72–2.25	1.22	0.59–0.79	0.66
N2	Flat small	0.73–1.96	1.32	0.51–0.72	0.61
N3	Round large	0.49–1.80	1.07	0.67–0.95	0.83
N4	Flat round	0.55–1.49	1.13	0.63–0.95	0.77
N5	Flat large	0.77–1.69	1.36	0.55–0.77	0.63
N6	Round small	0.56–1.86	1.09	0.59–0.89	0.73
N8	Mixture	0.55–1.78	1.17	0.54–0.83	0.65

The overall area was totaled for all frontal positions and all profile positions, and the difference was calculated. The difference of total projected areas in both positions were characteristically different for all fractions. For all flat fractions there were over 1,000 pixels, and for all round sets there were less than 1,000 pixels.

The coefficient of separation ranged from 1.22 to 1.36 for all flat sets and from 1.07 to 1.09 for all round sets. Circularity shape factors for all fractions differed consistently for the two orthogonal positions. The circularity shape factor was the best criterion to differentiate among fractions as a set separated by the mechanical separator. The classification based on pattern recognition was consistent with the separation by the mechanical classification device.

The results of this study illustrate and document the general approaches using image analysis and techniques that can be used to characterize cereal grains. The use of more sophisticated techniques, such as on-line computer software control, alone or in combination with other methods, to fully identify, grade, and characterize cereal grains in marketing channels and plant breeding programs is underway at the U.S. Grain Marketing Research Laboratory.

LITERATURE CITED

AHMED, N., and RAO, K. R. 1975. Page 225 in: Orthogonal Transforms for Digital Signal Processing. Springer-Verlag: New York.

ANDREWS, H. C. 1972. Introduction to Mathematical Techniques in Pattern Recognition. Wiley-Interscience: New York.

AMERICAN SOCIETY FOR TESTING AND MATERIALS. 1980. Annual Book of ASTM Standards, Designation D3849-80. The Society: Philadelphia, PA.

ATTLE, J. R., OUEY, D., and SWENSON, R. A. 1980. Application of image analysis. American Laboratory 4:85-99.

CHEN, C. H. 1973. Page 10 in: Statistical Pattern Recognition. Sparton/Hayden: New Jersey.

GONZALEZ, R., and WINTZ, P. 1977. Digital Image Processing. Addison-Wesley Publishing Co.: New York.

IMANCO. 1978. Instruction Manual—Quantimet 720. Imanco: Monsey, NY.

JURS, P. C., and ISENHOUR, T. L. 1975. Chemical Applications of Pattern Recognition. John Wiley and Sons: New York.

LOEB, H. 1984. A quantitative image analysis system. Am. Lab. 4: 55-65.

MARTIN, T. B., and REDDY, D. R., eds. 1975. IEEE Symp. Speech Recognition. IEEE Trans. Acoust. Speech Signal Process. 23(1):2-148.

POMERANZ, Y., CZUCHAJOWSKA, Z., MARTIN, C. R., and LAI, F. S. 1985. Determination of corn hardness by the Stenvert hardness tester. Cereal Chem. 62:108.

YOUNG, T. Y., and CALVERT, T. W. 1974. Classification, Estimation and Pattern Recognition. American Elsevier Publishing Co.: New York.

ZAYAS, I., POMERANZ, Y., and LAI, F. S. 1985. Pattern recognition for wheat variety discrimination. Cereal Foods World 30(8):563.

[Received March 1, 1985. Revision received July 22, 1985. Accepted July 23, 1985.]