
Dockage identification in wheat using machine vision

M. NAIR and D.S. JAYAS

Department of Biosystems Engineering, 438 Engineering Building, University of Manitoba, Winnipeg, MB, Canada R3T 5V6. Received 20 November 1997; accepted 15 October 1998.

Nair, M. and Jayas, D.S. 1998. **Dockage identification in wheat using machine vision.** *Can. Agric. Eng.* **40**:293-298. Algorithms were developed to classify dockage components from Canada Western Red Spring (CWRS) wheat and other cereal grains (e.g. durum wheat, barley, rye, and oats) based on morphological and colour features. The dockage classes used were: wheat heads, chaff, wild oats, canola, wild buckwheat, flax, and broken-wheat pieces. The wheat head dockage class was subdivided into single and multiple wheat heads; the subdivision improved the classification accuracy for wheat heads. The developed algorithms were tested on images taken with an area scan camera. Training and test data sets were established to evaluate the classification accuracies based on the extracted features. Morphology, colour, and morphology-colour models were evaluated for classifying the dockage components. Mean accuracies of 89.4% for the morphology model, 71.4% for the colour model, and 93.2% for morphology-colour model were obtained when tested on the independent test data sets using the holdout non-parametric classifier. **Keywords:** grain, classification, foreign material, automation, cleaning, grading, digital image processing, imaging.

Des algorithmes furent développés afin de classifier les types d'impuretés dans du blé rouge de printemps de l'ouest canadien et d'autres céréales (blé dur, orge, seigle et avoine) selon leurs caractéristiques morphologiques et leur couleur. Les catégories d'impuretés utilisées étaient: épi de blé, ivraie, folle-avoine, canola, sarrasin, lin et particules de grain de blé. On divisa la catégorie d'impureté "épi de blé" en deux: épi simple et épis multiples. Cette division permit d'augmenter la précision dans le classement des épis de blé. Les algorithmes furent testés avec des images obtenues à l'aide d'une caméra à balayage. Des ensembles de données pour le test et la formation furent créés pour évaluer la précision du classement fait à partir des éléments extraits des images. Les modèles morphologique, de la couleur et morphologie-couleur qui servent au classement des impuretés furent évalués. Lorsque les modèles furent testés avec un ensemble indépendant de données de test, la précision moyenne fut de 89.4% pour le modèle morphologique, de 71.4% pour le modèle de la couleur et de 93.2% pour le modèle morphologie-couleur. **Mots-clés:** céréales, triage, matières étrangères, automatisation, nettoyage, classement, traitement d'image numérisée.

INTRODUCTION

A machine vision system (MVS) could be used effectively for objective measurement of physical quality parameters of grain at terminal elevators (grain handling facilities). The primary reason for its potential application for wheat inspection at terminal elevators lies in its capability 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, colour, reflectance, and texture) (Sapirstein and Bushuk 1989).

Moreover, powerful microcomputers and specialized hardware have fostered moderately priced, high performance machine vision systems which are able to handle the wide variability in size, shape, colour, and textural characteristics of agricultural produce and products. As a result, the MVS offers the potential to improve the competitive position of the agri-food industry by raising product quality while lowering processing costs.

Fast and accurate information on the contents of a grain sample can be used to increase the efficiency of most grain handling operations (such as grain unloading, cleaning, binning, and shipping) at terminal elevators (Shatadal et al. 1995a, 1995b). The important applications of machine vision to the grain industry include the design and development of an objective, fast, and reliable on-line monitoring system for grain in continuous flow at many points in a terminal elevator. This would lead to increased cleaning throughput and enhanced recovery of salvageable grains. Use of machine vision guided controls and robotics could lead to complete automation of terminal elevators. A commercial MVS for grain inspection at terminal elevators is not yet available. Although substantial efforts have been made in the last decade on using MVS for automatic information acquisition on the content and quality of grain samples (e.g. Barker et al. 1992; Neuman et al. 1987; Shatadal et al. 1995a, 1995b; Symons and Fulcher 1988a, 1988b; Zayas et al. 1989, 1990), many of the special needs and problems involved in industrial application are still unresolved. Dockage identification is one of these needs.

At terminal elevators, grain is received in railcars and cleaned by mechanical separators before the grain is stored for shipping to export buyers. A series of cleaning machines are used to remove the dockage. Export shipments are prepared by blending grain from several storage bins of like type and grade to meet the buyers' specifications for quality and grade specifications. For the salvageable grain recovery and for adjusting the efficiency of the cleaning machines, dockage tester fractions have to be identified. With the present machine vision technology, the railcar contents can be identified and recognized as wheat, barley, durum, rye, and oats and the clean samples can be identified with reasonable accuracy (>95%). So far, no work has been reported on dockage identification in wheat using machine vision.

The objectives of this research were:

1. to use a machine vision system to identify dockage in wheat by developing algorithms to extract morphological and limited colour features from grain kernels and from dockage components,

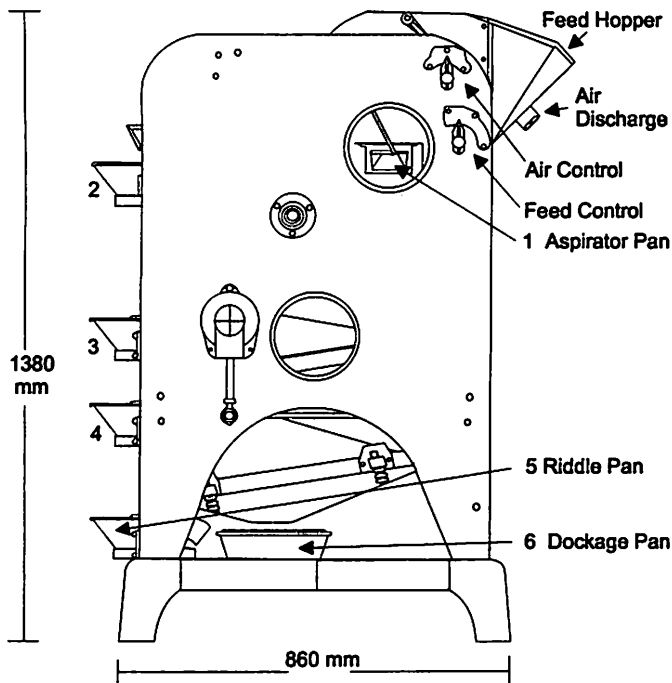


Fig. 1. Schematic of Carter dockage tester. The numbers 2, 3, and 4 are pan numbers.

2. to investigate the potential of different image-extracted morphological and colour features for classification of dockage components from CWRS wheat and from other cereal grains (e.g. durum wheat, barley, rye, and oats), and
3. to investigate the feasibility of classifying dockage components into their appropriate classes using the selected features by designing or selecting appropriate statistical pattern classifiers.

DOCKAGE IN WHEAT

Definition

Dockage is the material that is separable from the grain sample by the use of approved cleaning equipment in order that the grain can be assigned the highest grade for which it qualifies (Anonymous 1994). Dockage is reported in percentage by mass. The percentage of dockage in a sample is reported in increments of 0.5% when the grain is not commercially clean and in the export shipments to the nearest 0.1% (when authorized by the Canadian Grain Commission to contain dockage).

Composition of Dockage

Dockage is assessed by running the uncleaned representative sample of 500 or 1000 g through the Carter dockage tester (Fig. 1). As defined in the Grain Grading Handbook for Western Canada (Anonymous 1994), the following are the dockage constituents:

1. foreign material removed over the riddle, less any portion which is eligible for machine separation (pan 2, Fig. 1),
2. material removed by aspiration (pan 1, Fig. 1),
3. material removed by No. 5 buckwheat sieve in the lower position (pans 5 and 6, Fig. 1),

4. a maximum of 10.0 % by mass of soft earth pellets hand picked from the cleaned sample, and
5. any material removed by cleaning for grade improvement.

Dockage vs foreign material

Foreign material is defined as material other than grain of the same class, which remains in the sample after the removal of dockage. In wheat, other cereal grains like barley, rye, oats, and non-grain material like earth pellets, chaff, and fertilizer pellets are termed as foreign material if collected on pans 3 and 4 with wheat and are termed as dockage if they get separated by the dockage tester (Fig. 1).

Some of the dockage constituents (soybean, earth pellets, stones, and wild mustard) are not included in the dockage class as they are very rare constituents. Wheat heads, chaff, wild oats, flax, wild buckwheat, canola, and broken kernels are identified as dockage constituents. The main dockage constituents and other cereal grains (durum wheat, barley, rye, and oats) were used in this study to assess the capability of the morphological and limited colour features for their identification.

METHODS and MATERIALS

Vision hardware

The hardware of the image acquisition system used in this study is described in Shatadal et al. (1995a, 1995b) and Luo et al. (1997). The system consists of a colour camera (Model DXC-300A, Sony, Japan), a camera control unit (CCU) (Model M3, Sony, Japan), a diffuse illumination chamber with a circular fluorescent tube (305 mm diameter, 32 W, Model FC1279/CW, Philips, Singapore) and light controller (Model FX0648-2/120, Mercron, Richardson, TX), a colour monitor for on-line image display, a colour frame grabber installed in a personal computer, and an optical disk drive.

The camera captured images from the samples placed in the illumination chamber and outputted three parallel analog video signals, namely red (R), green (G), and blue (B), corresponding to the three NTSC (National Television System Committee) colour primaries, and a sync signal. The camera control unit performed the time-division multiplexing and dc restorations of the RGB signals and time signal generation for the frame grabber. The frame grabber digitized the RGB analog video signals to three 8-bit 512 x 512 size RGB digital images at a speed of 30 frames per second and stored them in three of the four on-board buffers. The acquired digital images were then transferred to the optical disk for storage to be retrieved later for analysis using a Pentium-75 personal computer.

Illumination standardization

A Kodak white card with 90% reflectance (E 152-7795, Eastman Kodak Co., Rochester, NY) was used as a white reference to standardize the illumination level. The lamp voltage was set to the rated value of 120 V. An image of the white card was acquired over a small central area of 50 x 50 pixels and the mean grey level values of R, G, and B bands were computed and used as illumination level indicators. By manually adjusting the iris control and performing the white balance with the CCU, all three values were adjusted to 250 ± 1 . For on-line monitoring, white balance may be done at the

Table I. Selection of features using STEPDISC analysis with thirteen classes in the model.

Number	Selected features	Average squared canonical correlation	Partial r^2
1	Length	0.0773	0.928
2	Average red	0.1590	0.980
3	Average green	0.2217	0.799
4	Haralick ratio	0.2842	0.760
5	Rectangular ratio	0.3399	0.686
6	First Fourier magnitude	0.3899	0.620
7	Standard deviation of radii	0.4058	0.285
8	Standard deviation of green	0.4258	0.336
9	Standard deviation of red	0.4406	0.267
10	Area ratio	0.4555	0.261
11	Maximum radius	0.4768	0.298
12	Second invariant moment	0.4847	0.221
13	First invariant moment	0.4935	0.309
14	Area	0.5007	0.158
15	Width	0.5055	0.101
16	Second Fourier magnitude	0.5112	0.094
17	Intensity	0.5142	0.068
18	Thinness ratio	0.5163	0.066
19	Fourth invariant moment	0.5182	0.045
20	Perimeter	0.5199	0.041
21	Third Fourier magnitude	0.5213	0.035
22	Minimum radius	0.5230	0.033
23	Radius ratio	0.5238	0.025
24	Fourth Fourier magnitude	0.5254	0.036
25	Standard deviation of blue	0.5259	0.023
26	Third invariant moment	0.5266	0.018
27	Length of minor axis	0.5268	0.011
28	Aspect ratio	0.5272	0.007
29	Length of principal axis	0.5273	0.002
30	Average blue	0.5278	0.001

beginning of an 8-h shift when using fluorescent tubes for illumination (Luo et al. 1997).

Grain and dockage samples

Composite grain samples of Canada Western Red Spring (CWRS) wheat (grades 1, 2, and 3), durum wheat (grades 1, 2, 3, and 4), barley (grades 1 and EX1), oats (grades 1 and 3), and rye (grade 1) were collected by the Industry Services Division of the Canadian Grain Commission, Winnipeg, MB, from different growing regions of western Canada for the 1994 growing season. From these samples, seven grain types (CWRS-1, CWRS-2, CWRS-3 wheat, durum wheat, barley, rye, and oats) from 20 growing regions were used. The regions were chosen using the climatic subdivisions of the Canadian

Prairies (Putnam and Putnam 1970). Three hundred kernels (25 kernels in an image frame) from each growing region were used for each grain type and grains from five randomly selected growing regions were analysed.

Dockage samples were obtained by running 15 kg of uncleaned samples of CWRS wheat from the Glenlea research farm of the University of Manitoba through the Carter dockage tester. One hundred grams of each dockage tester fractions were collected from the Industry Services Division of Canadian Grain Commission, Winnipeg, MB. The dockage components used for analysis were: wheat head, chaff, wild oats, canola, wild buckwheat, flax, and broken-wheat pieces. Based on the results of preliminary analysis, wheat head was partitioned into two classes: single- and multiple-wheat head (Nair 1997).

Image acquisition

The system was stabilized for 30 min. The illumination standardization and white balancing were done and repeated after every three images. Black background was used for all samples except for samples of canola, wild buckwheat, wild oats, and flax where white background was used for better thresholding. For on-line monitoring, however, only a single background colour is preferable with improved algorithms for thresholding. We hope to develop such algorithms in the future. In each frame 25 objects were placed, imaged, and stored as digital images on the optical disk for further analysis.

Pattern classification

After converting the rectangular pixel images into square pixel images (Luo et al. 1997), the images were thresholded using automatic thresholding (Parker 1994). Holes were filled and extra regions were deleted from the thresholded image. Morphological and basic colour features were extracted from the labelled and original images, respectively. The extracted features are listed in Table I with their level of contribution to the classifier.

Discriminant analyses using PROC DISCRIM of SAS (1990) were carried out using cross-validation (leave-one-out), and holdout methods. In each case, normal and non-parametric estimations were used. In the non-parametric estimation, the k-nearest neighbour method was used with a k value of 5. In the holdout method, cereal grains from four randomly selected growing regions (300 kernels per growing region) were used as the training data set and from one growing region as the test data set. In the cross-validation method, the training data set used in the holdout method was used for classification.

To determine the level of contribution by individual features for classification, PROC STEPDISC (SAS 1990) was used. The training data set used in the holdout method was used for feature selection in STEPDISC analysis. Individual rankings of features were determined using STEPDISC analysis by removing the best feature from the model and by re-ranking the remaining features. For example, in a model with five features

Table II. Summary of classification accuracies (%) for different models and different analysis methods.

Model	Analysis method	CWRS♣	Durum⊗	Barley	Rye	Oats	Mwheat⊛	Chaff	Wild oats	Canola	Wbwheat ♠	Flax	Broken ♣	S-wheat §	Mean
Morphology-colour 12†	Leave-one-out	99.2	96.6	98.2	96.3	99.8	96.3	92.6	99.8	99.4	98.5	99.3	97.6	¶	97.8
	Normal														
	Leave-one-out	99.8	99.1	98.3	97.2	99.9	98.9	97.8	99.8	99.5	99.7	99.8	98.8	-	99.1
	Non-par ♣														
	Holdout	99.7	89.7	95.3	99.0	99.7	27.3	30.0	99.3	99.7	98.7	99.3	98.0	-	86.3
Morphology 13	Normal														
	Leave-one-out	100.0	95.1	95.9	91.8	100.0	98.2	93.6	99.3	98.9	97.3	99.0	85.4	97.2	96.3
	Non-par														
	Holdout	100.0	83.0	92.3	95.7	99.3	98.3	8.7	48.7	98.7	95.3	96.7	87.3	95.7	84.6
	Normal														
Colour -13	Holdout	100.0	79.3	97.3	96.0	99.7	100.0	19.0	94.3	99.0	98.0	98.3	87.3	94.0	89.4
	Non-par														
	Leave-one-out	65.9	65.1	81.3	74.1	31.3	58.3	75.9	60.3	95.1	96.7	51.9	85.8	56.3	69.0
	Normal														
	Leave-one-out	78.5	77.4	84.3	85.7	75.4	72.2	81.9	94.2	84.0	92.8	95.9	78.6	56.1	81.3
Morphology-colour-13	Non-par														
	Holdout	77.7	74.7	65.0	70.0	0.0	71.7	34.0	46.3	96.3	99.0	52.3	82.7	55.7	63.5
	Normal														
	Holdout	71.6	88.0	51.0	75.0	2.7	81.7	50.0	94.7	83.7	96.7	95.3	78.3	59.0	71.4
	Non-par														
Morphology-colour-13	Leave-one-out	99.3	94.7	97.6	95.8	99.8	96.7	83.3	99.6	99.4	98.2	99.3	97.3	96.8	96.8
	Normal														
	Leave-one-out	100.0	98.8	98.6	97.3	100.0	99.0	96.8	99.8	99.6	99.7	99.8	98.8	98.9	99.0
	Non-par														
	Holdout	99.9	89.7	95.0	98.7	99.3	99.3	9.7	99.0	99.7	98.7	99.3	98.0	95.0	90.9
Morphology-colour-13	Normal														
	Holdout	100.0	96.7	98.3	99.3	99.3	100.0	21.7	99.7	100.0	100.0	100.0	99.7	96.7	93.2

♣ Canada Western Red Spring Wheat, ⊗ Durum Wheat, ⊛ Multiple-wheat head, ♠ Wild buckwheat, ♣ Broken-wheat pieces, § Single-wheat head, † Number of classes in the model, ¶ Single-Wheat heads were included in the multiple-wheat head class, ♣ Non parametric estimation.

the STEPDISC analysis was carried out with four features (the best feature from the five features model was removed) and the four features were ranked. This process was repeated with one feature in the final model. The classification accuracies with the best 5, 10, and 15 features were determined and compared. Morphology, colour, and morphology-colour models were developed and compared by using the mean classification accuracies.

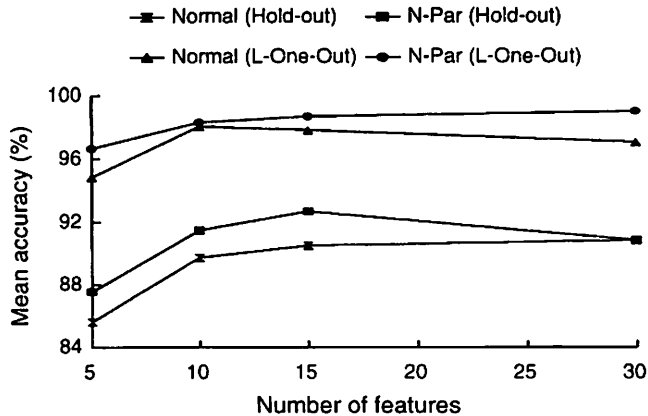


Fig. 2. Comparison of classification accuracies with selected features.

RESULTS and DISCUSSIONS

Morphology and colour model classifiers

Classification accuracies for different classes and the mean accuracies obtained using four classifiers (leave-one-out normal, leave-one-out non-parametric, holdout normal, and holdout non-parametric) are summarised in Table II for four models (Morphology-colour-12, Morphology-13, Colour-13, and Morphology-colour-13). The numbers 12 and 13 in the models represent the number of classes. The mean classification accuracies obtained using the leave-one-out method were higher than those obtained using the holdout method (Table II). Because the training set is used for classification in the leave-one-out method, higher classification accuracies are expected. The lower mean values for the holdout method were mainly caused by large classification errors for the chaff class. For example, the classification errors for the chaff using the leave-one out method for Morphology-colour-12 model were 92.6 to 97.8% whereas using the holdout method were 12 to 30% (Table II). For the Morphology-colour-12 model, the classification accuracies were also low (2.3 to 27.3%) for the wheat head class using the holdout method. For other classes, the classification accuracies were similar for both methods with only a few exceptions. These low classification accuracies were caused by the large range in the morphological features for these two classes. The wheat head class had many single wheat heads (i.e. a single kernel inside the husk) and multiple wheat heads (i.e. two or more kernels inside the husk and attached together). Therefore, the wheat head class was divided into single and multiple wheat head classes, resulting in 13 classes. No such division was possible for the chaff class.

The holdout method classifier is suitable for industrial application because a classifier can be developed prior to

implementation for testing and classifying objects on-line. Therefore, the classification accuracies obtained using the holdout non-parametric classifier for different models with 13 classes are discussed in the following paragraphs.

The classification accuracies for the model with only the morphological features for the holdout method with the non-parametric estimation were: CWRS wheat (100.0%), durum wheat (79.3%), barley (97.3%), rye (96.0%), oats (99.7%), multiple-wheat head (100.0%), chaff (19.0%), wild oats (94.3%), canola (99.0%), wild buckwheat (98.0%), flax (98.3%), broken-wheat pieces (87.3%), and single-wheat head (94.0%).

The classification accuracies for the model with only colour features for the holdout method with the non-parametric estimation were: CWRS wheat (71.6%), durum wheat (88.0%), barley (51.0%), rye (75.0%), oats (2.7%), multiple-wheat head (81.7%), chaff (50.0%), wild oats (94.7%), canola (83.7%), wild buckwheat (96.7%), flax (95.3%), broken-wheat pieces (78.3%), and single-wheat head (59.0%) (Table II). The classification accuracies for the colour model were poor. A model with only basic colour features is not helpful in discriminating the dockage components from wheat. A model containing morphological and colour features was developed and evaluated.

Morphology-colour model classifier

The classification accuracies for the morphology-colour model with thirteen classes for the holdout method with non-parametric estimation were: CWRS wheat (100.0%), durum wheat (96.7%), barley (98.3%), rye (99.3%), oats (99.3%), multiple-wheat head (100.0%), chaff (21.7%), wild oats (99.7%), canola (100.0%), wild buckwheat (100.0%), flax (100.0%), broken-wheat pieces (99.7%), and single-wheat head (96.7%) (Table II). The division of the wheat head class into two classes (single and multiple) improved the classification accuracy significantly (from <30% to >95%, Table II).

The classification accuracies for chaff were low. Chaff was mis-classified as oats because of the closeness of their features to the oats class. Inclusion of textural and additional colour features may improve the classification accuracy for chaff. All dockage components except chaff could be classified with >95% accuracy by a machine vision system. Chaff should be removed by aspiration if machine vision is to be used for optimizing a cleaning unit.

Selection of features

The features were arranged in the descending order of their level of contribution to the classifier (Table I). Length was the most significant [average squared canonical correlation (ASCC) = 0.0773] and the average blue was the least significant (ASCC = 0.5278) feature used in the model with thirteen classes. Discriminant analyses were carried out with the first 5, 10, and 15 features from Table I and the classification accuracies were compared with all 30 features (Fig. 2).

The classification accuracies were low when the first five features were used. The mean classification accuracy increased with the number of features up to the first 15 features and remained constant thereafter except for the holdout method with non-parametric analysis where it decreased slightly (Fig. 2). The addition of more features did not improve the

performance of the classifier. It is important to note that the mean classification accuracy for the holdout method with all 30 features in the model with 13 classes was around 90% although 100% classification was achieved in many classes (CWRS wheat, canola, wild buckwheat, and flax). This was because the classification accuracy for chaff class was low.

SUMMARY and CONCLUSIONS

For determination of physical quality of wheat samples by machine vision, dockage constituents have to be identified and classified. In this study, discrimination capabilities of morphological and basic colour features were evaluated for identification of dockage constituents from CWRS wheat and other cereal grains. Morphology, colour, and morphology-colour models were developed and compared.

Based on this study, the following conclusions were made:

1. The model with only colour features was not sufficient to discriminate the dockage components from wheat (mean classification accuracy was about 70%).
2. The morphology model with 13 classes discriminated the dockage components with 84 to 96% mean accuracies whereas the morphology-colour model with 13 classes discriminated with 90 to 99% mean accuracies.
3. It is necessary to improve the classification accuracy of chaff for practical implementation.

ACKNOWLEDGMENTS

We thank Prince Pupert Grain Ltd., Natural Sciences and Engineering Research Council of Canada, and Agriculture and Agri-Food Canada for partial funding of this study. We thank the Industry Services Division of the Canadian Grain Commission, Winnipeg, MB, for assisting in the collection of the grain samples.

REFERENCES

Anonymous. 1994. *Official Grain Grading Guide*. Winnipeg, MB: Canadian Grain Commission.

Barker, D.A., T.A. Vouri, M.R. Hegedus and D.G. Myers. 1992. The use of ray parameters for the discrimination of Australian wheat varieties. *Plant Varieties and Seeds* 5:35-45.

Luo, X., D.S. Jayas, T.G. Crowe and N.R. Bulley. 1997. Evaluation of light sources for machine vision. *Canadian Agricultural Engineering* 39:309-315.

Nair, M. 1997. Dockage identification in wheat using machine vision. Unpublished M.Sc. thesis. Department of Biosystems Engineering, University of Manitoba, Winnipeg, MB.

Neuman, M., H.D. Sapirstein, E. Shweddyk and W. Bushuk. 1987. Discrimination of wheat class and variety by digital image analysis of whole grain samples. *Journal of Cereal Science* 6:125-132.

Parker, J.R. 1994. *Practical Computer Vision using C*. New York, NY: John Wiley and Sons, Inc.

Putnam, D.F. and R.G. Putnam. 1970. *Canada: A Regional Analysis*. Toronto, ON: J.M. Dent and Sons, Inc.

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.

Shatadal, P., D.S. Jayas and N.R. Bulley. 1995a. Digital image analysis for software separation and classification of touching grains: I. Disconnect algorithm. *Transactions of the ASAE* 38:635-643.

Shatadal, P., D.S. Jayas and N.R. Bulley. 1995b. Digital image analysis for software separation and classification of touching grains: II. Classification. *Transactions of the ASAE* 38:645-649.

Symons, S.J. and R.G. Fulcher. 1988a. Determination of wheat kernel morphological variation by digital image analysis: I. Variation in Eastern Canadian milling quality wheats. *Journal of Cereal Science* 8:211-218.

Symons, S.J. and R.G. Fulcher. 1988b. Determination of wheat kernel morphological variation by digital image analysis: II. Variation in cultivars of soft white winter wheats. *Journal of Cereal Science* 8:219-229.

Zayas, I., H. Converse and J. Steele. 1990. Discrimination of whole from broken corn kernels with image analysis. *Transactions of the ASAE* 33:1642-1646.

Zayas, I., Y. Pomeranz and F.S. Lai. 1989. Discrimination of wheat and non-wheat components in grain samples by image analysis. *Cereal Chemistry* 66:233-237.