

Detecting Vitreous Wheat Kernels Using Reflectance and Transmittance Image Analysis

Feng Xie,¹ Tom Pearson,^{2,3} Floyd E. Dowell,³ and Naiqian Zhang¹

ABSTRACT

Cereal Chem. 81(5):594–597

The proportion of vitreous durum kernels in a sample is an important grading attribute in assessing the quality of durum wheat. The current standard method of determining wheat vitreousness is performed by visual inspection, which can be tedious and subjective. The objective of this study was to evaluate an automated machine-vision inspection system to detect wheat vitreousness using reflectance and transmittance images. Two subclasses of durum wheat were investigated in this study: hard and vitreous of amber color (HVAC) and not hard and vitreous of amber color (NHVAC). A total of 4,907 kernels in the calibration set and 4,407 kernels in the validation set were imaged using a Cervitex 1625 grain inspection system. Classification models were developed with stepwise discriminant

analysis and an artificial neural network (ANN). A discriminant model correctly classified 94.9% of the HVAC and 91.0% of the NHVAC in the calibration set, and 92.4% of the HVAC and 92.7% of the NHVAC in the validation set. The classification results using the ANN were not as good as with the discriminant methods, but the ANN only used features from reflectance images. Among all the kernels, mottled kernels were the most difficult to classify. Both reflectance and transmittance images were helpful in classification. In conclusion, the Cervitex 1625 automated vision-based wheat quality inspection system may provide the grain industry with a rapid, objective, and accurate method to determine the vitreousness of durum wheat.

Vitreousness is an important international grading attribute in assessing the quality of durum wheat. Higher vitreousness indicates higher protein content, a harder kernel, coarser granulation, higher yield of semolina, superior pasta color, improved cooking quality, and opportunity for premium sales pricing (Dowell 2000). Vitreous kernels are glassy and translucent, while nonvitreous kernels are chalky and opaque. Some minor defects such as bleached, cracked, or checked hard vitreous kernels are considered vitreous. In contrast, sprouted kernels, foreign materials, scabby kernels, etc., are considered nonvitreous.

The current standard method of evaluating vitreousness of durum wheat in the United States is by manually inspecting a 15-g sample that is free of shrunken and broken kernels (USDA 1997). Because it is a subjective method, inspectors may disagree on classification. Therefore, an objective, automated, reproducible, and rapid method for determining durum wheat vitreousness is needed. Such a grading method should greatly reduce grain inspectors' subjectivity and labor, and benefit wheat producers, processors, and handlers.

Researchers have investigated various methods for evaluating durum wheat vitreousness. A perfect match with inspector classifications of obviously vitreous or nonvitreous durum wheat using near-infrared spectroscopy (NIRS) was reported by Dowell (2000). Classification rates of 91.1–97.1% were reported when studying dark, hard, vitreous, and nonvitreous hard red spring wheat using NIRS (Wang et al 2002). A Perten 4100 single-kernel characterization system was applied to detect wheat vitreousness (Sissons et al 2000; Nielsen et al 2003). This method was faster than the NIRS method, but classification capability was limited. Sissons et al (2000) reported ≈25–35% error of prediction with SKCS.

Machine vision techniques have been used to determine grain quality and classify grain cultivars based on color and geometry features (Zayas et al 1996; Ruan et al 1998; Luo et al 1999; Majumdar and Jayas 2002a,b). Image analysis has also been applied to study wheat vitreousness (Symons et al 2003; Wang et al 2003).

Wang et al (2003) developed an artificial neural network (ANN) model using reflectance images captured by a real-time image-based grain-inspection machine (Foss GrainCheck-310). Classification rates of vitreous and nonvitreous subclasses were 85–90% (Wang et al 2003). Symons et al (2003) found significant agreement between inspector-determined and machine-determined percentages of hard vitreous kernels in samples. Transmittance images of individual kernels were imaged with a monochrome camera.

It was hypothesized that using both reflectance and transmittance images would improve classification rate of vitreousness. The objective of this study was to evaluate an automated machine vision inspection system for detecting wheat vitreousness using both reflectance and transmittance images captured by a new grain inspection system.

MATERIALS AND METHODS

Sample Preparation

The Grain Inspection, Packers, and Stockyards Administration (GIPSA) of the United States Department of Agriculture collected samples for this study. Durum wheat samples were classified into two subclasses: hard and vitreous of amber color (HVAC) and not hard and vitreous of amber color (NHVAC). For this study, the HVAC subclass was further separated into three categories and NHVAC into six categories (Table I). Samples were visually reinspected by the Board of Appeals and Review (BAR) to check correct segregation. Samples in each category were divided into

TABLE I
Summary of Sample Categories and Sizes Used in This Study

Subclass and Category	Calibration	Validation	Reproducibility
HVAC ^a			
Clean	695	507	80
Bleached	507	507	40
Cracked	507	507	40
NHVAC ^b			
Clean	507	507	...
Bleached	468	429	...
Mottled	741	468	40
Sprouted	507	468	...
Foreign materials	468	507	...
All other categories	507	507	...
Total	4,907	4,407	200

^a Hard and vitreous of amber color durum wheat.

^b Not hard and vitreous of amber color durum wheat.

¹ Biological and Agricultural Engineering Department, Kansas State University, 147 Seaton Hall, Manhattan, KS 66506.

² Corresponding author. E-mail: tpearson@gmprc.ksu.edu

³ USDA-ARS, Grain Marketing and Production Research Center, 1515 College Avenue, Manhattan, KS 66502.

*The e-Xtra logo stands for "electronic extra" and indicates that the online version contains a color version of Fig. 1 not included in the print edition.

two sets of 1,000 kernels each: a calibration set and a validation set. The calibration set was used to develop classification models, and the validation set was used to test the models.

Image Data Collection

An automated vision-based wheat quality inspection system (Cervitec 1625 Grain Inspector, Foss Tecator, Hoganas, Sweden) was used to collect images. The machine is capable of continuously collecting either reflectance or transmittance images of 1,000 single kernels in <2 min. To collect reflectance and transmittance images of the same kernel, the feeder hopper on the wheel was removed so that kernels would stay on the wheel until removed by hand. Images were transferred to a PC and saved after capture. For reflectance image collection, white light was used as a light source and color reflectance images were collected. For transmittance images, red light was used as a light source because very little blue and green light would transmit through the kernel due to scattering. Figure 1 shows examples of reflectance and transmittance images. Each image was ≈ 110 pixels long by 50 pixels tall for a total of $\approx 5,500$ pixels per image. The kernels filled $\approx 80\%$ of the image area for a total of $\approx 4,400$ pixels. Reflectance and transmittance images of 4,907 calibration samples were collected. To test model performance, 4,407 validation samples were measured by following the same procedure (Table I). To test model classification reproducibility, 80 clean-HVAC kernels, 40 bleached-HVAC kernels, 40 cracked-HVAC kernels, and 40 mottled-NHVAC kernels were randomly selected from the calibration sample set and used as testing samples. Samples were imaged by the Cervitec 1625 system following the same procedure. The test was repeated six times. Samples were visually reinspected by five FGIS BAR inspectors after mechanical analysis.

Image Feature Extraction and Classification Procedure

Cumulative histograms of pixel intensities were computed for each color component of the reflectance image and for the red component of the transmittance image only. Two histograms for each image were computed: one consisted of raw counts of kernel pixels with an intensity between a specified range; and the second histogram consisted of the percentage of total kernel pixels with an intensity between a specified range. Five intensity values were used for each histogram bin, starting at zero all the way up to 255, for a total of 51 histogram bins for each image. The kernel portion of each reflectance image was segmented from the background by simple thresholding using software supplied by Foss. Transmittance images were segmented after the image was transferred from the Cervitec 1625 to the PC. Before thresholding, the transmittance image was first minimum filtered using a 3×3 pixel window (Jain 1989). If a pixel in the minimum filtered image had an intensity value >20 , it was considered to belong to a kernel.

Histogram features of both reflectance and transmittance images from the same kernel were saved together after extraction. Both raw histograms and percentage histograms were used. Each histogram contained 51 bin values and the entire 51 element vector of each histogram were used as potential discriminating features. Stepwise discriminant analysis (SAS Institute, Cary, NC) was applied to select important feature variables for classification models and determine the best variable combination for classification. The variable significance to enter or leave the model was 0.01. Discriminant models were developed using the best variable combination and then applied to validation samples and samples used for the reproducibility test.

An artificial neural network (Matlab, The MathWorks, Novi, MI) was also used to classify kernels as HVAC or NHVAC. Only reflectance images were used to build ANN models. Features from the red, green, and blue (RGB) color space images were used, as well as features from hue, saturation, and intensity (HSI) color space images. A total of 12 features from the two image color spaces were extracted. These features included the means of

blue, green, and red pixel intensities of each RGB image; the means of the saturation and intensity of all kernel pixels in each HSI image; means of the $\cos(\text{Hue})$ and $\sin(\text{Hue})$ pixel distributions; and histogram bin values from the red pixel intensity distribution, saturation distribution, hue distribution, intensity distribution, and normalized distribution of the intensity component of the HSI image. Values of the input nodes were projected into an input vector that contained these 12 extracted features from each image. The ANN models were developed using 150 epochs in the first run and using the number of epochs suggested by Matlab in the second run. The number of hidden neurons and layers was 10 and 1, respectively, for all models. ANN models were applied to validation samples, and results were compared with those of the other models.

The most effective model was selected based on the classification accuracy rates of each subclass. Classification rate for a subclass is equal to the number of kernels correctly classified as this subclass divided by the total number of kernels of this subclass in a sample.

RESULTS AND DISCUSSION

Two models were developed using the reflectance and transmittance images. The two-subclasses model classified kernels as either HVAC or NHVAC. The all-categories model grouped kernels into nine HVAC and NHVAC categories. For both models, the classification accuracy was simply the percentage of kernels correctly classified as HVAC or NHVAC. To calculate the classification accuracy for the all-categories model, all the kernels of each category correctly classified as HVAC or NHVAC were added up, then divided by the total number of HVAC or NHVAC kernels. When 10 features were selected by stepwise discriminant analysis for these two models, the all-categories model could correctly determine 94.1% of all HVAC and 85.6% of all NHVAC, while the two-subclasses model could correctly identify 89.4% of all HVAC and 82.4% of all NHVAC. These results show that the all-categories model was more accurate than the two-subclasses model in predicting wheat vitreousness. Most likely, the all-categories model is more accurate than the two-subclasses model because it takes into account the variance of each category, while the two-subclasses model lumps the variances of all categories together. Separating the variances allows the classifier to have more local, well-defined, centroids in the feature space for classifying the kernels.

The all-categories model was optimized using the SAS program. Both pooled and nonpooled covariance matrices were used to compute the discriminant functions. The most effective model was nonpooled matrices and equal prior probabilities of any subclass with 17 features selected. Feature variables were selected from both reflectance or transmittance images, including 11 intensity values of red, blue, or green reflectance images; two intensity values of transmittance images; and four histogram binary values of transmittance images. Classification rates for HVAC and NHVAC were 94.9 and 91.0%, respectively, for the calibration set and 92.4 and 92.7%, respectively, for the validation set (Table II). The best model developed by Wang et al (2003) was a two-classes model of weighted sample sizes. It classified 87.6% HVAC and

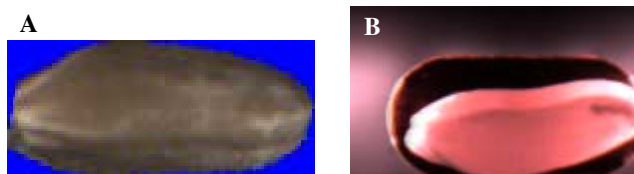


Fig. 1. A, Example of reflectance images. B, Example of transmittance images.

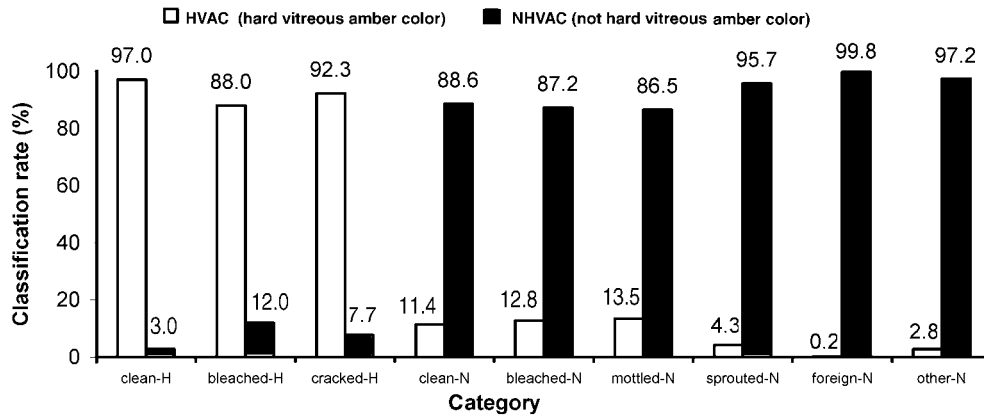


Fig. 2. Classification rates for all HVAC and NHVAC categories predicted by the all-categories discriminant model.

TABLE II
Classification Rates (%) of Discriminant and ANN Models for HVAC and NHVAC in Calibration and Validation Sample Sets

Description	Calibration			Validation		
	HVAC	NHVAC	Average	HVAC	NHVAC	Average
Discriminant models						
All-categories ^a	94.9	91.0	92.9	92.4	92.7	92.6
Reflectance images only	92.0	79.0	85.5	89.2	82.5	85.8
Transmittance images only	89.2	85.4	87.3	87.0	90.5	88.7
ANN models						
ANN150 all-categories ^b	92.6	92.8	92.7	85.1	90.6	87.9
ANN82 all categories ^c	92.3	92.0	92.1	83.6	89.6	86.6
ANN150 two-subclasses ^d	94.7	95.0	94.9	86.0	91.5	88.8
ANN73 two-subclasses ^e	96.0	94.1	95.1	88.7	92.7	90.7

^a Using both reflectance and transmittance images.

^b Using 150 epochs.

^c Using 82 epochs.

^d Using 150 epochs.

^e Using 73 epochs.

TABLE III
HVAC and NHVAC (%) in the Composition Sample Predicted by Cervitec 1625 System and FGIS BAR Inspectors

Test Number ^a	Cervitec 1625		Inspectors	
	HVAC	NHVAC	HVAC	NHVAC
1	80	20	70.6	29.4
2	78	22	75.3	24.7
3	80.5	19.5	71.6	28.4
4	80	20	78.4	21.6
5	79.5	20.5	77.6	22.4
6	81	19
Average	79.8	20.2	74.7	25.3
SD	1	1	3.5	3.5
Max	81	22	78.4	29.4
Min	78	19	70.6	21.6

^a Samples measured using the Cervitec 1625 system six times (tests 1–6) and visually inspected by five inspectors (tests 1–5).

91.6% NHVAC kernels correctly. Thus, the all-categories model developed in this study is likely more accurate in determining wheat vitreousness than the two-classes model (Wang et al 2003).

Figure 2 shows classification rates for each category predicted by the all-categories model. The all-categories model correctly classified 97% clean-HVAC while only 88% bleached-HVAC were determined correctly. Bleached-NHVAC had the second lowest classification accuracy among all other categories, showing that bleached kernels were hard to segregate. Mottled-NHVAC had the lowest classification accuracy among all the other categories. Wang et al (2003) also met the same problem. These kernels are characterized as having mottled or chalky spots randomly located inside the kernels. Even though use of transmittance image increased classification accuracy, results showed that mottled kernels were the most difficult category to determine by the all-categories

model. Collecting reflectance or transmittance images from the opposite side of the kernel only increased accuracy 1–2% (data not shown). Feature variables used in the all-categories model were selected from either reflectance or transmittance images, which indicates that the use of both reflectance and transmittance images were helpful for classifications.

The all-categories model classification reproducibility was tested (Table III). The testing sample set consisted of 80% HVAC and 20% NHVAC. The all-categories model predicted HVAC and NHVAC percentages in the sample set. The all-categories model had an average of 79.8% HVAC with a standard deviation of 1.0%. The FGIS BAR inspectors had an average of 74.7% HVAC with a standard deviation of 3.5%. These results show that the Cervitec 1625 system could predict wheat vitreousness more accurately and precisely than visual inspection.

TABLE IV
Classification Accuracy (%) for Four Categories in the Composition Sample

Test Number	Clean-HVAC	Bleached-HVAC	Cracked-HVAC	Mottled-NHVAC
1	98.8	92.5	85.0	75.0
2	98.8	85.0	80.0	72.5
3	98.8	90.0	87.5	72.5
4	95.0	92.5	82.5	65.0
5	98.8	82.5	87.5	70.0
6	98.8	82.5	87.5	70.0
Average	98.1	87.5	85.0	70.8
SD	1.5	4.7	3.2	3.4

Classification rate for each category in each run is presented in Table IV. Averaged classification rate for clean-HVAC, bleached-HVAC, cracked-HVAC, and mottled-NHVAC was 98.1, 87.5, 85.0, and 70.8%, respectively. Clean-HVAC had a higher rate than that in the validation set, which was 97.0% (Fig. 2). Clean-HVAC had the smallest standard deviation (1.5%) of all the categories, showing that clean-HVAC was the easiest category to segregate. The standard deviation for bleached-HVAC was the highest (4.7%), showing that bleached-HVAC was hard to determine for the all-categories model. Mottled-NHVAC had the lowest classification rate and the second highest standard deviation (3.4%), confirming that mottled-NHVAC was difficult to classify. This might be due to random positioning of mottled spots inside the kernels.

Models using either reflectance or transmittance images alone were developed to classify all categories (Table II). Classification rates for HVAC and NHVAC of these two models did not produce results as good as those obtained from the all-categories model using both imaging modes. For the validation set samples, the reflectance model could classify 89.2% of all HVAC kernels and 82.5% of all NHVAC kernels correctly. The transmittance model could determine 87.0% of all HVAC and 90.5% of all NHVAC correctly. The reflectance model had higher classification accuracy for HVAC, while the transmittance model had higher classification accuracy for NHVAC, indicating that reflectance images were more favorable for HVAC determination, while transmittance images were more helpful for NHVAC classification. Results confirmed that model accuracy was improved by use of both reflectance and transmittance images.

Features extracted from reflectance images were entered into the Matlab ANN software, which suggested using 73 epochs instead of 150 epochs when training the two-subclasses ANN model. It recommended using 82 epochs instead of 150 epochs when training the all-categories ANN model. All these models were applied to the validation set samples (Table II). The two-subclasses ANN model using 73 epochs was the most accurate ANN model. It correctly classified 88.7% of all HVAC kernels and 92.7% of all NHVAC kernels in the validation set. This model had higher classification accuracy than the all-categories discriminant model when training the models using calibration samples. But results were not as good as those of the all-categories discriminant model after application to the validation set, which was 92.4 and 92.7% for HVAC and NHVAC, respectively. It is important to note that only reflectance images were used during ANN training because ANN cannot handle the transmittance images output from the Cervitec 1625 system, which may be the reason for lower classification accuracies compared with the all-categories discriminant model that used features from both transmittance and reflectance models.

Features from both reflectance and transmittance images were input into another ANN software program, the Neuroshell Classifier (Ward Systems Groups, Frederick, MD), to build a two-subclasses ANN model. Results using this software were similar to those from discriminant analysis. Therefore, use of ANN or discriminant analysis should not make a large difference on model performance.

CONCLUSIONS

A discriminant model trained and tested on a validation set classified 92.4% of HVAC kernels and 92.7% of NHVAC kernels correctly. Standard deviation for predicting HVAC and NHVAC percentage in a composition sample was 1%, whereas the SD of trained inspectors was 3.5%. Mottled-NHVAC and bleached-HVAC kernels were difficult to classify. Developing this model required 17 histogram features from both reflectance and transmittance images. Results showed that an automated, vision-based, wheat quality inspection system, such as the Cervitac 1625, has potential as a rapid, objective, and accurate method for determining wheat vitreousness that could be more accurate than visual inspection.

ACKNOWLEDGEMENTS

We thank the Grain Inspection, Packers, and Stockyards Administration of United States Department of Agriculture and Foss North America for investing in this research. We also thank Amanda Gleason for help with data collection.

LITERATURE CITED

- Dowell, F. E. 2000. Detecting vitreous and nonvitreous durum wheat kernels using near-infrared spectroscopy. *Cereal Chem.* 77:155-158.
- Jain, A. K. 1989. *Fundamentals of Digital Image Processing*. Prentice Hall: Englewood Cliffs, NJ.
- Luo, X., Jayas, D. S., and Symons, S. J. 1999. Comparison of statistical and neural network methods for classifying cereal grains using machine vision. *Trans. ASAE* 42:413-419.
- Majumdar, S., and Jayas, D. S. 2002a. Classification of cereal grains using machine vision. II. Color models. *Trans. ASAE* 43:1677-1680.
- Majumdar, S., and Jayas, D. S. 2002b. Classification of cereal grains using machine vision. III. Texture models. *Trans. ASAE* 43:1681-1687.
- Nielsen, J. P., Pedersen, D. K., and Munck, L. 2003. Development of non-destructive screening methods for single kernel characterization of wheat. *Cereal Chem.* 80:274-280.
- Ruan, R., Ning, S., Song, A., Ning, A., Jones, R., and Chen, P. 1998. Estimation of fusarium scab in wheat using machine vision and a neural network. *Cereal Chem.* 75:455-459.
- Sissons, M. J., Osborne, B. G., Hare, R. A., Sissons, S. A., and Jackson, R. 2000. Application of the single-kernel characterization system to durum wheat testing and quality prediction. *Cereal Chem.* 77:4-10.
- Symons, S. J., Schepdael, L. V., and Dexter, J. E. 2003. Measurement of hard vitreous kernels in durum wheat by machine vision. *Cereal Chem.* 80:511-517.
- USDA. 1997. *Grain Inspection Handbook. II. Grading Procedures*. Grain Inspection, Packers and Stockyard Administration: Washington, DC.
- Wang, D., Dowell, F. E., and Dempster, R. 2002. Determining vitreous subclasses of hard red spring wheat using visible/near-infrared spectroscopy. *Cereal Chem.* 79:418-422.
- Wang, N., Dowell, F. E., and Zhang, N. 2003. Determining wheat vitreousness using image processing and a neural network. *Trans. ASAE* 46:1443-1150.
- Zayas, I. Y., Martin, C. R., Steele, J. L., and Katsvech, A. 1996. Wheat classification using image analysis and crush-force parameters. *Trans. ASAE* 39:2199-2204.

[Received November 17, 2003. Accepted March 15, 2004.]