Going against the Grain: Real-Time Classification of Grain Quality

Zachary Pezzementi[1], Carl Wellington[1], Trenton Tabor[1], Cason Male[1], Herman Herman[1], and Scott Miller[2]

***Abstract*— In this work, we present a vision and learning-based approach for detecting and estimating the composition of grain flowing through a combine harvester. It is implemented as a module in John Deere's Auto Maintain function, which uses this information to perform closed loop control of machine settings to maintain operator-set performance goals. We describe the hardware and software design, performance evaluation against different sources of ground truth information, and development process to produce a commercial product.**

I. INTRODUCTION

Modern combine harvesters, or simply combines, are highly advanced and complex mobile factories. Through a series of carefully orchestrated mechanical operations, these machines efficiently reap ripe crop, removing it from the field; thresh the crop, removing grain from the outer husk; and clean the grain, separating it from chaff (left over husk material).

In order to maximize combine effectiveness it is necessary for an operator to configure the machine's settings to optimize for a number of factors. These include:

1) Minimizing grain loss (clean grain ejected back to the field)
2) Removing the hull from grain without breaking it
3) Separating the grain from the rest of the plant and any other foreign material

Starting settings are typically provided by the combine manufacturer or an expert operator. In an ideal harvest 100% of available grain kernels would be perfectly threshed, cleaned, and deposited into the grain tank with all other materials exiting the machine back to the field. Unfortunately, this is not feasible when attempting to maintain speed of harvest and fuel efficiency.

In addition to configuring machine settings, the operator is responsible for numerous primary tasks, including driving, system monitoring, and grain tank unloading on the go. When field conditions change, machine settings are often no longer optimal. This occurs regularly throughout a daily harvesting operation. Correcting machine settings for changing conditions can be very difficult, as a result of the complex interactions between the various internal mechanisms of the harvester. In many situations, even an expert operator cannot compensate for geo-temporal conditions such as wet spots or wind bursts, occurrences that can be short-lived but high-frequency in nature. Additionally, the operator's information is limited, as they can only directly view a highly biased sample through a small window into the grain tank.

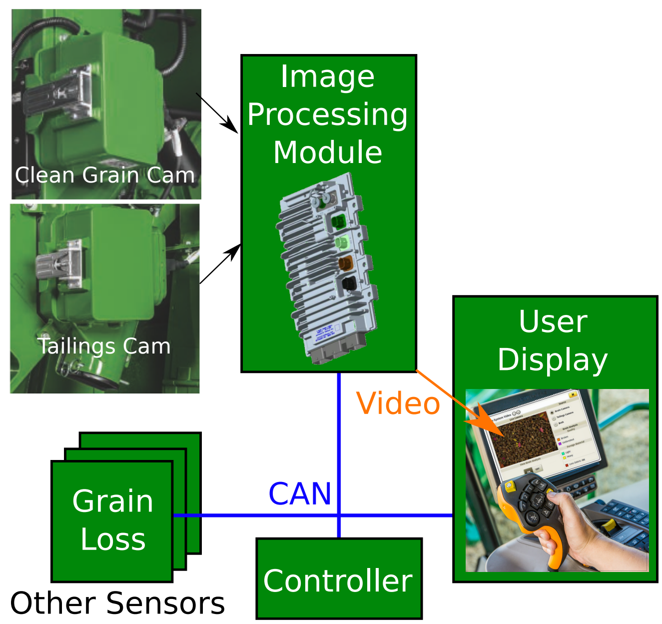

Fig. 1: System diagram, showing how the image processing module, grain loss module, and other sensors provide inputs to the controller, which closes the loop to maintain performance.

A combine operating at peak harvesting performance can provide financial benefit to the farmer. Providing consistency in the harvested crop, such as minimizing broken grain kernels or foreign material present in the harvested grain, can result in improved results when the crop is sold [1]. In a typical farming operation of several thousand acres, these benefits can represent savings as high as hundreds of dollars per hour. [2].

A system to automate settings optimization can allow machines to operate with much greater effectiveness, reduce operator load, and enable even novice operators to maintain expert operator levels of harvesting performance. Ultimately, such automation results in higher profitability for the individual farmer and more food available to feed the world. To realize such a system, we apply a vision and learning approach to estimate the quality of the grain [3]; this is then combined with other sensors and a controller to enable automated adjustment and optimization of machine settings, using the design shown in Figure 1. The goal of the controller is to maintain the balance of speed, loss, and quality established by an operator's choice of initial settings as conditions

*This work was sponsored by John Deere.

[1]National Robotics Engineering Center, Carnegie Mellon University Robotics Institute, Pittsburgh, PA 15201, USA.

[2]John Deere Intelligent Solutions Group, Urbandale, IA 50322, USA `MillerScott@JohnDeere.com`

(a) Corn (b) Wheat

Fig. 2: Example images of corn and wheat, demonstrating variety of appearance within a single crop.

change, much like cruise control in an automobile. The full system, John Deere Integrated Combine Adjustment - ICA 2, was shown at Agritechnica, received an award, and was recognized for its improved automatic adjustment of combine settings, such as threshing settings [4].

There has been significant work in visual analysis of grain characteristics for automated grading [5], classifying different varieties [6], and sorting based on quality [7]. The survey in [8] provides more references, but this prior work is generally geared towards controlled lab or processing settings. The focus of this work is to perform grain quality estimation on an active combine to allow closed loop optimization of the machine performance. One other system performs visual analysis of quality on a combine for display to an operator, but does not use it for control [9], [10].

To perform live analysis of the grain as it is being harvested, we image it as it moves through the machine. Then we estimate how much of each image is new material, whether there are obscurants present in the image, and how much of each class of interest is present. In this work, we focus on the final part of the pipeline: quality estimation. We treat it as a pixelwise classification problem, for which we define five standard classes of interest, and try to discriminate between them:

- Clean Grain - Whole particles of grain, with hulls removed. Material farmers want to maximize.
- Broken Grain - Grain that has been chipped or broken into smaller pieces.
- Unthreshed Grain - Grain that is still in its hull. (Does not apply to corn or canola.)
- Material Other than Grain (MOG)
  - MOG-Light - Material that can be blown away by a fan. Includes leaves, broken hulls, and other thin plant material.
  - MOG-Heavy - Material that is too dense to blow away with a fan. Includes sticks and, in the case of corn, cobs.

This classification is performed for our five supported crops: wheat, corn, soy, canola, and barley. More details of our algorithmic approach follow in Section III.

It is important to note that there is great appearance variety within each of these five crops. Grain has been bred all around the world and exposed to different environmental conditions. It is impractical to know the exact crop variety that a combine will harvest a priori. Our approach must therefore be robust to a wide variety of appearance within each crop, as shown in Figure 2.

II. HARDWARE

Effective image-based measurement of the material flowing through a combine begins with capturing high quality images that contain representative quantities of the classes of interest. This is challenging because the material moves at high speeds, there is significant dust, and the physics of how the three-dimensional flow of material moves through the system can cause biases. As shown in Figure 1, the combine includes a small bypass off the main elevator containing the final cleaned grain. This bypass was originally designed to collect samples of the grain for measuring moisture; we reuse the bypass to mount a camera (see Figure 3), which captures a sequence of images of representative samples of the material being collected, to measure its different constituents. As shown in Figure 1, we mount a second camera to the tailings elevator to make similar measurements of the grain in the tailings.

Each camera has a small glass window that provides a view of the material. Strobe lighting inside the camera is synchronized to the image capture so we can capture a blur-free image even when the material is moving. We also optimized the lighting to improve image classification by ensuring even lighting across the viewing window, minimizing internal reflections.

The images from the two cameras are passed to an embedded processing unit containing a single core processor with a limited amount of memory, so all steps of the image pipeline must be efficient to maintain fast update rates and ensure that all the material flowing can be processed.

III. QUALITY ESTIMATION

The pipeline for processing images and producing weight composition estimates is depicted in Figure 4. Multiple features of two different types (patch and segment) are extracted from images. A feature sampler extracts feature vectors from a regular grid of sampling locations, and these feature vectors are then passed to a classifier, which produces class confidence estimates for each of these image samples.

Fig. 3: Photo of camera hardware (right) mounted on clean grain bypass (left). It is shown in open position (to allow cleaning and maintenance), which allows views into both the bypass and the camera. During operation, it would be closed, with camera glass flush against the bypass window.

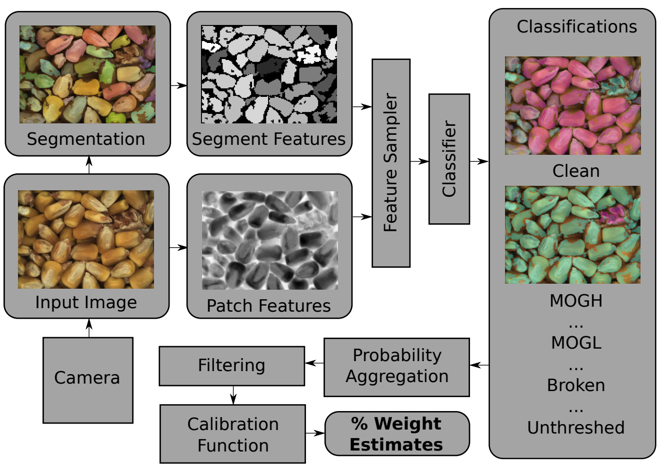

Fig. 4: Classification pipeline, showing each step of the process from feature extraction to producing weight composition estimates for an image.

Then a filtering and calibration process assembles many per-sample estimates into weight composition estimates. Each step is described in more detail below.

A. Feature extraction

We use two different types of features for classification: one set that operates on patches of the image, and another that operates on regions from a segmentation process.

Patch features extract color and texture information for square regions of the image at different scales. Color features capture average color in each channel of the YCbCr color space, and normalized versions of individual components for improved brightness invariance. Texture features are computed by applying the FFT or DCT to a patch of the image and scaling the result down to a 4×4 grid of bins. These were applied using square patches starting at size 4×4 pixels and increasing in size by powers of 2.

Segment features are extracted from pre-processed segmented images. Each image undergoes a watershed segmentation [11] process to separate each piece of grain and plant material from one another and from the image background. Segment level features are extracted from each resulting foreground segment in the image. Many features are geometric in nature, such as contour length, roundness, convexity, concavity and major/minor axes. Non-spatial segment features include segment average color, min/max brightness, intensity and variances of color within the segment.

The features associated with a given pixel in the image depend on the patches and the segment that it falls into. The Feature Sampler assembles a feature vector from all features associated with each chosen pixel's segment and all features from each patch (across different scales) the pixel lies within.

We began with a large and diverse set of feature extractors that originally produced hundreds of features per pixel from this process. We then implemented a feature selection technique to reduce this to a robust subset, maintaining performance while reducing processing time to enable real-time operation on our embedded processor.

In the first stage of feature selection, aimed at cutting a significant portion of less-important features, pruning was applied to large blocks of related features. Feature blocks ranged in size from tens to roughly a hundred features of a similar feature family, whose computation was tied together. A feature family, for example, could include a type of related linear filters. A set of experiments was run to decide which feature block to prune by iteratively removing feature blocks and evaluating the effects on precision and recall. (More details on this performance evaluation follow in Section IV.) Blocks that proved to have little impact on performance were removed from the final feature set.

We followed this by a second, finer feature selection step, where we assembled a feature set by iteratively adding features one at a time. In each step, the feature that produced the largest performance improvement was added to the set. After this process, we reduced our feature set of several hundred features to roughly half its original size, meeting our computational requirements.

B. Classification

Once a set of features are assembled for a sampled image location, they are passed to a classifier to estimate class likelihoods at that location. To generate training data, a team of experts hand-labeled the classes of interest in a large set of representative image data collected across different varieties and conditions.

We experimented with several different options for the type of classifier used in order to investigate the trade-off between accuracy and computation. As in the feature selection experiments, performance was quantified based on the precision/recall curve on a controlled validation set, described further in Section IV. We evaluated different structures for combining a set of very fast binary SVM classifiers, but we found superior performance with acceptable runtime from a single, non-linear, multiclass classifier using a multi-layer perceptron [12]. We use a standard classifier network containing a single hidden layer with rectified linear units [13] feeding into a final softmax classification, trained with stochastic gradient descent [14] to keep training time and memory footprint manageable. We tuned the hyperparameters for this classifier by sampling from a Gaussian process model of the classifier performance as a function of these hyperparameters. This allowed us to efficiently sample the hyperparameter space based on performance and uncertainty to quickly hone in on a high performing network with a minimal number of time-consuming re-training and evaluation iterations.

C. Filtering and Calibration

Classifier outputs are interpreted as estimates of probability of the presence of each class for every pixel sampled. These probability estimates are aggregated across all samples for an image to get an estimate of the class distribution within the image. A visualization of the resulting probability map is shown in Figure 4 for a sampling scheme that samples at regular grid points across the image.

Since only a small amount of material is visible in each image, evidence must be accumulated over several images to get stable estimates of the composition of material moving through the combine. This can be seen through the noisy variation in the per-image estimates shown as dots in Figure 5. A windowed averaging filter produces a smoothly-varying signal suitable for use in closed-loop control, as shown by the solid line in the figure. The system accounts for the amount of new image visible at each time step to prevent counting the same material multiple times.

Finally, a calibration function is applied to the filtered outputs to convert them from a distribution of image samples to percent by weight compositions, the desired output of the system. This is accomplished through a regression to ground truth weight samples described further in Section IV and shown in Figure 7.

IV. PERFORMANCE ANALYSIS AND TUNING

The performance of the classification step can be readily evaluated with precision-recall (PR) curves, such as the ones in Figure 6. Modifications to the classifier configuration, such as changes to the feature set, training data, or training parameters, were evaluated in terms of area under the PR curve. The multi-class nature of the problem was handled through a weighted averaging of the areas for the classes of interest. This allowed us to produce a single number for comparison of various possible configurations. Then, once a particular configuration was chosen, the same plots guided selection of the desired trade-off between precision and recall for each class.

The filtering step was tuned through the generation of time plots like the one shown in Figure 5. Since relatively little material is seen in each image, and typical compositions of

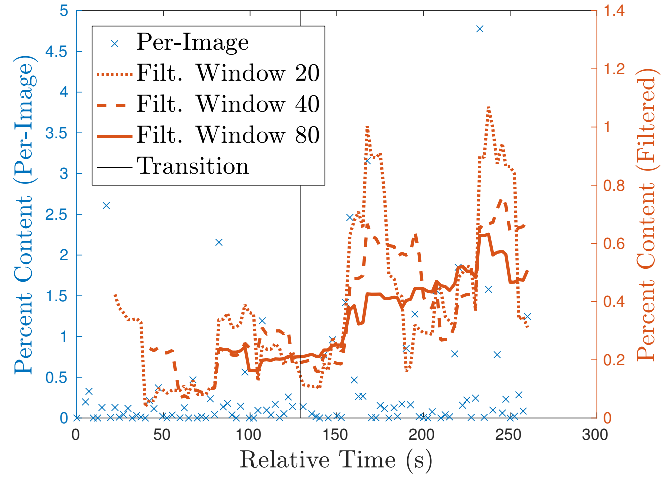

Fig. 5: Sample time plot for a transition from low (0.1%) to high (0.6%) amount of cobs, showing both per-image content estimates and the result of three different windowed averaging filters. Note that even in a higher cobs content condition, most images have little to no cobs visible, so filtering is important.

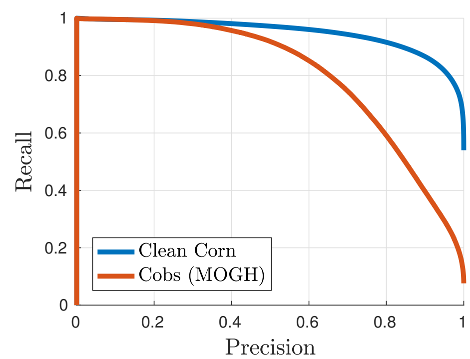

Fig. 6: Sample per-pixel PR curves for detection of clean corn and cobs respectively.

interest are on the order of a few percent by weight, effective filtering of per-image results is essential. Evaluating a diverse set of plots for different magnitude changes we wished to detect allowed tuning the filter configuration to trade off between noise suppression and lag in detecting changes.

Finally, system-level performance can be evaluated by comparing classifier outputs for a set of images to ground truth percent-by-weight compositions for the material associated with those images. These ground truth measurements were collected using a Carter-Day mechanical granular materials separator, the same machine used by the U.S. Department of Agriculture and Canadian Grain Standards Commission. The process employed is identical to that used by grain elevators where farmers sell harvested grain [15], [16]. Figure 7 shows several such comparisons, for samples with a range of different ground truth compositions of cobs. A regression between these points is used to determine the calibration function between image samples and weight percentages. In this case, a simple linear regression describes

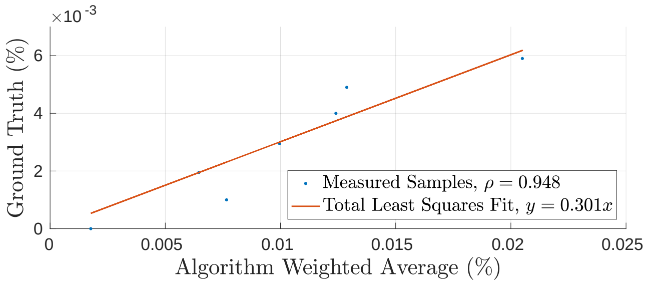

Fig. 7: Samples-to-weight calibration plot, showing correlation between algorithm outputs and ground truth weight percentages.

the relationship well, with a bivariate correlation of 0.948. The strength of this fit can then also be used to evaluate the system-level accuracy.

V. CONCLUSIONS

We have presented a system for estimating the composition of a material flow, applied to a combine harvester for monitoring quality in harvested grain and performing closed-loop control to maintain performance. We described the development process of this system, including hardware design, the algorithmic approach based on classical vision and machine learning, and how we optimized and tuned the system to meet computational and performance constraints for commercial deployment.

These same techniques could also be applied to other problems estimating the composition of bulk material. Potential applications include later stages of grain processing, conveyor belt systems in industrial and mining applications, recycling, and other biological and medical image analysis.

REFERENCES

[1] United States Department of Agriculture, *7CFR8.810 Official United States standards for grain*, 2018. [Online]. Available: https://www.gipsa.usda.gov/fgis/gradingprocedures.aspx

[2] K. Lawton, "Smarter combine technology for 2018," *Corn and Soybean Digest*, June 2017. [Online]. Available: http://www.cornandsoybeandigest.com/equipment/smarter-combine-technology-2018

[3] C. K. Wellington, A. J. Bruns, V. S. Sierra, J. J. Phelan, J. M. Hageman, C. Dima, H. Boesch, H. Herman, Z. A. Pezzementi, C. R. Male, J. Campoy, and C. Vallespi-Gonzalez, "Grain quality monitoring," U.S. Patent US2016/0 189 007 A1, June 30, 2016.

[4] Agritechnica, "The silver medals," *Agritechnica Innovations Magazine*, 2015. [Online]. Available: https://www.agritechnica.com/fileadmin/downloads/2015/neuheiten/AT_Neuheitenmagazin_2015_eng_IT.pdf

[5] L. A. I. Pabamalie and H. L. Premaratne, "A grain quality classification system," in *2010 International Conference on Information Society*, June 2010, pp. 56–61.

[6] N. S. Visen, J. Paliwal, D. Jayas, and N. White, "Image analysis of bulk grain samples using neural networks," in *2003 ASAE Annual Meeting*. American Society of Agricultural and Biological Engineers, 2003.

[7] P. Brckner, K. Anding, H. Weiensee, and M. Dambon, "Quality assurance of grain with colour line scan cameras," in *Proc. OPTO SENSOR+TEST*, Nrnberg, Germany, 2009, pp. 139–144.

[8] P. Vithu and J. Moses, "Machine vision system for food grain quality evaluation: A review," *Trends in Food Science & Technology*, vol. 56, pp. 13 – 20, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S092422441630084X

[9] W. Behnke, "Device for detection and determination of the composition of bulk material," U.S. Patent US2012/0 004 815 A1, January 28, 2014.

[10] Agritechnica, "The silver medals," *Agritechnica Innovations Magazine*, 2013. [Online]. Available: https://www.agritechnica.com/fileadmin/_migrated/content_uploads/Agritechnica_2013_Innovations-Magazine.pdf

[11] J. Serra, *Image analysis and mathematical morphology*. Academic Press, Inc., 1983.

[12] C. M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2006.

[13] V. Nair and G. E. Hinton, "Rectified linear units improve restricted boltzmann machines," in *Proceedings of the 27th international conference on machine learning (ICML-10)*, 2010, pp. 807–814.

[14] Y. Lecun, L. Bottou, G. B. Orr, and K.-R. Müller, "Efficient backprop," in *Neural networks: tricks of the trade*, G. B. Orr and K.-R. Müller, Eds. Springer, 2003, ch. 1, pp. 9–50.

[15] United States Department of Agriculture, *Equipment Handbook*, August 2016. [Online]. Available: https://www.gipsa.usda.gov/fgis/handbook/equipmentHB/EquipmentHB_2016-08-04.pdf

[16] Canadian Grain Commission, *Official Grain Grading Guide*, August 2017. [Online]. Available: http://www.grainscanada.gc.ca/oggg-gocg/oggg-gocg-2017-eng.pdf