

Modeling and Calibrating Visual Yield Estimates in Vineyards

Stephen Nuske, Kamal Gupta, Srinivasa Narasimhan and Sanjiv Singh

Abstract Accurate yield estimates are of great value to vineyard growers to make informed management decisions such as crop thinning, shoot thinning, irrigation and nutrient delivery, preparing for harvest and planning for market. Current methods are labor intensive because they involve destructive hand sampling and are practically too sparse to capture spatial variability in large vineyard blocks. Here we report on an approach to predict vineyard yield automatically and non-destructively using images collected from vehicles driving along vineyard rows. Computer vision algorithms are applied to detect grape berries in images that have been registered together to generate high-resolution estimates. We propose an underlying model relating image measurements to harvest yield and study practical approaches to calibrate the two. We report on results on datasets of several hundred vines collected both early and in the middle of the growing season. We find that it is possible to estimate yield to within 4% using calibration data from prior harvest data and 3% using calibration data from destructive hand samples at the time of imaging.

1 Introduction

Harvest yield prediction is critical to any vineyard grower for deciding when and how to make adjustments to their vines to optimize growth, for preparing a grower for the harvest operation, for shipping their crop, storing their crop and also selling their crop on the market. The typical process of estimating yield is for workers to manually sample a small percentage of the vineyard and extrapolate these mea-

Stephen Nuske, Srinivasa Narasimhan and Sanjiv Singh
Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, U.S.A., 15213

Kamal Gupta
Indian Institute of Technology Delhi, India

Corresponding author: Stephen Nuske, e-mail: nuske@cmu.edu

surements to the entire vineyard. The manual practice is labor intensive, expensive, inaccurate, spatially sparse, destructive and riddled with subjective inputs.



(a) Chardonnay vineyard used in experiments



(b) Sensors used in experiments



(c) Raw images

(d) Detected berries

Fig. 1: Photos of the vineyards and equipment from our experiments. Equipment mounted on an aluminum frame fixed on the back tray of a Kawasaki Mule farm utility vehicle. The sensing equipment used is a Nikon D300s color camera facing sideways from the vehicle detecting the fruit, an AlienBees ARB800 ring flash mounted around the lens of the color camera illuminating the scene, a PointGrey BumbleBee2 stereo camera facing back down the row tracking the vehicle motion. Images collected of the vines are processed to detect and measure the vine fruit.

Vineyard managers currently do not have the information they require to make informed decisions on their operations with accuracy and precision. We present technology that can make dense predictions of harvest yield efficiently and automatically using cameras. Here we report results of an approach to automatically detect and count grapes to forecast yield with both precision and accuracy. The approach is to take conventional visible light cameras through a vineyard to image the vines and detect the crop and predict yield, see Fig. 1b.

Visually detecting grape berries is difficult because often there is a lack of color contrast to the background, which is often similarly colored to the grapes, (Fig. 1a). We specifically address the issues of lighting and lack of color contrast, by using shape and texture cues for detection. Also not all berries are visible and therefore the camera makes incomplete measurements that need to be calibrated.

Preliminary results of our approach were reported in Nuske et al. [9], where we presented a method to detect berries and count berries to correlate against yield and we extend our prior work in the following ways:

1. we study the underlying model relating the image measurements of the grape berries to the harvest weight
2. we demonstrate two different approaches to calibrate our image berry measurements to harvest yield
3. we present experimental results with image data collected at various stages during the growing season and in both wine and table-grape vineyards

We deployed our method on three different vine varieties and conducted experiments in which manual per-vine harvest weights were collected and used as ground truth to evaluate our automated yield measurements. The size of the experiment includes 860 individual vines, of three different varieties and vineyards over a linear distance of 2.5km. Our method predicts weight with approximately 4% error of the overall actual harvest yield and approximately 8% error of the harvest weight for individual vineyard rows.

2 Related Work

Current practices to forecast yield are inaccurate because of sampling approaches that tend to adjust towards historical yields and include subjective inputs (Clingel-effer et al. [2]). The calculation of final cluster weight from weights at véraison use fixed multipliers from historic measurements, Wolpert and Vilas [12]. Unfortunately, multipliers are biased towards healthier vines thus discriminating against missing or weak vines and multipliers for cluster weights vary widely by vineyard, season and variety.

Sensor-based yield estimation in vineyards has been attempted with trellis tension monitors, multispectral sensors, terahertz-wave imaging and visible-light image processing. A dynamic yield estimation system based on trellis tension monitors has been demonstrated (Blom and Tarara [1]) but it requires permanent infrastruc-

ture to be installed. Information obtained from multispectral images has been used to forecast yields with good results but is limited to vineyards with uniformity requirements (Martinez-Casasnovas and Bordes [6]). A proof of concept study by Federici et al. [5] has shown that terahertz imaging can detect the curved surfaces of grapes and also has the potential to detect these through occluding thin canopy. The challenge for this approach is to achieve fast scan rates to be able to deploy the scanner on a mobile platform.

Small scale yield estimation based on simple image color discrimination has been developed by Dunn and Martin [4]. This approach was attempted on Shiraz post-véraison (i.e. after color development, very close to harvest) in short row segments. The method would not be applicable for the majority of real world examples where the fruit appears over a background of similarly-colored leaves, as is the case in white grape varieties and in all varieties before véraison. Other recent small scale experiments in vineyard work is Dey [3] present a method for classifying plant structures, such as the fruit, leaves, shoots based on 3D reconstructions generated from image sequences which unlike our work is sensitive to slight wind whilst imaging. Other crop detection based on computer vision methods using color pixel classification or shape analysis has been attempted on various fruit types – Jimenez et al. [7] provides a summary of fruit detection work, Singh et al. [10] present a method for detecting and classifying fruit in apple orchards and Swanson et al. [11] use the shading on the curved surfaces of oranges as a cue for detection.

Our prior work, Nuske et al. [9], demonstrated how to detect berries similar in color to the background of leaves using a combination of image processing techniques including a radial symmetry transform and classification of texture properties. In the previous section we list the extensions to our earlier work.

3 Modelling Harvest Yield with Image-based Measurements

Viticulturalists have long studied the process of predicting the size of the harvest yield and have developed models of the various yield components (Clingeffer et al. [2]). In the most basic form, the weight of the harvest (W_h) can be expressed as a product of the number of berries (N_b) and the mean weight of the berries (W_b).

$$W_h = N_b W_b \quad (1)$$

Once berry set has passed the number of berries can be considered constant throughout the remainder of the season, whereas the mean berry weight will substantially increase in size. The standard practice for generating accurate yield predictions is to measure the number of berries early in the season and use historic information of the mean berry weight to project to a harvest estimate. Using historical records of mean berry weight will introduce error into the projection as the mean berry weight at harvest will vary from year to year. However the variation in berry weight is small and is known only to account for 10% of the year-to-year variations

in harvest yield. The berry count contributes to the remaining 90% (Clingeffer et al. [2]) of year-to-year yield variations.

Our approach to predict the yield at harvest time is also focussed on measuring the number of berries and we do so using cameras mounted on vehicles driving collecting images of the vines. From a single input image we use our visual berry detection algorithm detailed in Nuske et al. [9] for calculation of the number of berries. Often the canopy of the vines visually similar to the fruit. To detect the fruit we find possible berry locations based on their shape for robustness and classify based on texture and color. The method is to look for radially symmetric points in a fixed, pre-decided radius range followed by color and texture based classification. Fig. 1d presents examples of the berries being detected by our algorithm.

The number of berries found by our visual berry detection algorithm (N_b^d) is the measurement that we pass to a yield forecasting function, $f(\cdot)$, which outputs an estimate (\hat{N}_b) of the actual berry count:

$$\hat{N}_b = f(N_b^d) \quad (2)$$

Multiplying our estimated berry count by the expected berry weight at harvest (W_b) gives our yield prediction.

To generate accurate estimates, the function f must model several biases that are inherent to the visual detection process. In Nuske et al. [9] we treated the biases together as a single first order linear factor (k):

$$\hat{N}_b = f(N_b^d) = \frac{1}{k} N_b^d \quad (3)$$

Here in this work we study the individual causes of the bias in an attempt to better understand the system as a whole. First, we introduce three different types of occluders that cause many berries to not be visible to the camera and bias the counts:

1. Self-occlusions (k_s): Berries hidden behind berries within the same grape cluster
2. Cluster-occlusions (k_c): Berries hidden behind other grape clusters
3. Vine-occlusions (k_v): Berries hidden behind the leaves and shoots of the vine

All of these physical biases can be seen in Fig. 1. There are also biases in the visual detection process. In Nuske et al. [9] the performance of the detection algorithm was analyzed to find that a fraction of the berries visible to the camera are not detected by the algorithm and also small fraction of times the algorithm falsely reports a berry where there was not, and these two factors combined introduces a detection bias (k_d). Further, when combining detections from several overlapping images, the system is susceptible to errors where berries are either double-counted or mistakenly not counted, introducing a mis-registration bias, k_r .

The naive approach is to combine these bias terms as linear factors as follows:

$$f(N_b^d) = \frac{N_b^d}{k_s k_c k_v k_d k_r} \quad (4)$$

Later, in the results section, we attempt to isolate the five bias parameters aiming to deepen our understanding of the origins of the error in our visual estimation framework.

3.1 Estimating Self-occlusions

In our prior work (Nuske et. al [9]) we used the visible berry count as a prediction of the cluster size, assuming the visible berry count is proportional to the total berry count:

$$N_b \propto N_b^d \quad (5)$$

In the results section we study the visible berry count in controlled experiments and also in an attempt to improve the measurement of the occluded berries in a cluster, we propose two potential modifications.

The first alternative measurement we propose is to take the convex hull formed by all the berries in the cluster. Assuming the cluster has uniform density, k , and an average thickness of the grape cluster to be d , we multiply the area A to this fixed cluster depth, and normalize with the average berry radius R_b .

$$N_b \propto d \frac{A}{R_b^2} \quad (6)$$

The second alternative to measuring the size of a cluster is to extend the convex hull model by predicting the berries occluded by the visible layer of berries using an ellipsoid model. A grape cluster's volume can be approximated with an ellipsoid cutting off the image plane as an ellipse. We find the best fit ellipse for the berry center locations with same normalized second central moments. Given the semi-axes of the ellipse in pixels R_1 and R_2 , with $R_1 \geq R_2$, volume of the corresponding ellipsoid would be proportional to volume occupied by the berries (B_v) in the cluster. Using the average berry radius (R_b) of the cluster, we can hence calculate the total number of berries occupied by the cluster:

$$\begin{aligned} V_c &\propto \frac{4}{3} \pi r_1 r_2^2 \\ N_b &= V_c / \left(\frac{4}{3} \pi r_b^3 \right) \end{aligned} \quad (7)$$

We study these three approaches to measuring grape cluster size in controlled laboratory tests in the results section. In the vineyard experiments we focus just on using the visible berry count will evaluate the alternatives in future vineyard experiments once we have developed a method to accurately segment and separate neighboring clusters.

4 Calibrating Image Measurements to Harvest Yield

Here we take into consideration different procedures that are possible for calibrating our image measurements. In our prior work (Nuske et al. [9]) we demonstrated that computing a ratio between berries detected and the harvested fruit on one portion of data is sufficient for predicting yield on another portion of data by the applying a linear ratio to the image measurement. In practice, knowing the mean occlusion ratio of a given vineyard would be needed at the time of imaging, because it would defeat the purpose of predicting yield if it were necessary to wait for the harvest data before it were possible to measure the occlusion ratio. We propose two methods for acquiring the calibration ratio at (or before) the time of imaging, well in advance of harvest.

4.1 Calibration of occlusion ratio from destructive hand samples at time of imaging

One approach takes a small number of destructive hand samples in the vineyard at the time of imaging. The vines are imaged first and then on a small sample of vines the fruit is destructively removed and weighed. The hand fruit weight is projected to harvest using the ratio between current berry weight and expected berry weight at harvest. Taking the hand estimate against the image berry count for these specific vines produces an occlusion ratio that can be estimated well in advance of harvest, and applied to predict yield of the remaining vines that were not destructively sampled.

Fig. 2a shows a satellite image of the Chardonnay vineyard, highlighted with red to indicate the six vineyard rows that were imaged in our experiment. On bottom row, purple marks indicate the 15 vines in which the hand samples used for calibration. In Fig. 2b a graph shows the relationship between the hand fruit samples collected the day after imaging and the image berry counts. We derive a calibration function from this relationship and predict the crop weight based on the image berry counts of the remaining vines that were not a part of the destructive hand sample.

4.2 Calibration of occlusion ratio from prior year harvest data

The second method we evaluate for calibrating the image measurements is to use harvest data from prior growing seasons. We have analyzed harvest data from vines trained and prepared in a similar manner from year to year and noticed consistencies. The advantage of calibrating from a prior harvest season is that hand samples are not necessary. The method simply takes a total measurement of the fruit harvested and compared against the berry count detected in the imagery.

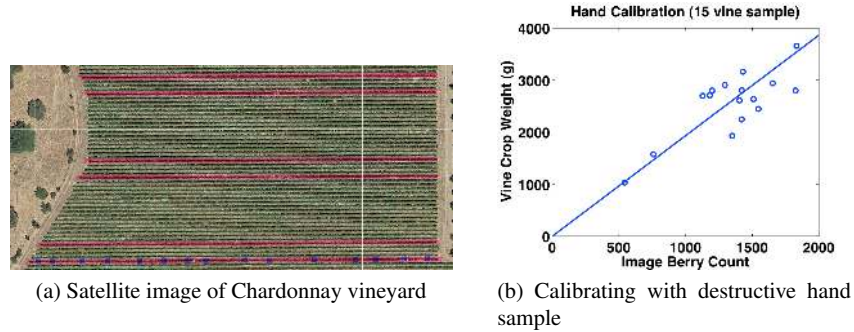


Fig. 2: Satellite image of the Chardonnay vineyard. Highlighted with red to indicate rows that were imaged by our setup and marked with purple on the bottom row to indicate where destructive hand samples were measured the day after imaging. Overall six rows were imaged totaling 665 vines, and 15 vines on the bottom row were destructively hand sampled. Graph showing calibration between the hand sample and the corresponding image berry counts.

We compute calibration functions from datasets collected in 2010 and normalize the calibration based on the mean berry weight at harvest. We then apply the calibration to the image berry counts we collected a year later in 2011 in a vineyard of a different varietal. In the following results section we compare the accuracy of the various approaches to predict harvest yield.

5 Results

5.1 Vineyard Datasets

We deployed our method on three different vine varieties and conducted experiments in which manual per-vine harvest weights were collected and used as ground truth to evaluate our automated yield measurements. The size of the experiment is significant, including 860 individual vines, totaling 2.5km vines, including following varieties: Traminette, Riesling and Chardonnay. See Table 1 for details of the different datasets, and see Fig. 1b for an example of the equipment we use to image the vines. Equipment mounted on an aluminum frame fixed on the back tray of a Kawasaki Mule farm utility vehicle. The sensing equipment used is a Nikon D300s color camera facing sideways from the vehicle detecting the fruit, an Alien-Bees ARB800 ring flash mounted around the lens of the color camera illuminating the scene, a PointGrey BumbleBee2 stereo camera facing back down the row tracking the vehicle motion and a synchronization box generating pulses to keep the two

cameras synchronized. We also collected a small dataset in controlled laboratory for studying self-occlusions, as described in the following section.

To generate a measure for each vine we use the stereo camera using a visual odometry algorithm (Kitt et al. [8]) to track position along the row. We reduce drift in the odometry by detecting the stakes supporting the trellis infrastructure which are fixed at known spacings and provide an extra source of positioning. We detect the stakes by simple image processing searching for near vertical lines in the images. Finding the stakes in neighboring frames enable us to triangulate the stakes location and depth from the camera. We assume the fruiting zone lies in a single plane at the depth of the stakes. Using this assumption we can compute the overlapping content of the images and crop images to avoid double counting the fruit.

Table 1: Vineyard Dataset Description

Variety	Location	Date	Trellis	Time before harvest at imaging	Mean berry weight at imaging	Num. vines
Traminette	Fredonia, NY	Sep. 2010	VSP	10 days	1.6g	88
Riesling	Fredonia, NY	Sep. 2010	VSP	10 days	1.5g	124
Chardonnay	Modesto, CA	June 2011	Semi-VSP	90 days	0.15g	648

5.2 Evaluating Berry Self-occlusion

First, we evaluate the occlusion of berries within a cluster by the outer layer of clusters (k_b) and study some approaches to potentially improve the estimate of the number of hidden berries. For this specific study, we use a controlled laboratory environment where we collected images individually of 56 grape clusters. We use ripe clusters of the Thompson Seedless variety. For each cluster we collected several images from different orientations, at a fixed distance, and collected a weight and a count of the number of berries. In the laboratory dataset we do not use our automatic detection algorithm and instead hand mark all berries visible within the images to replicate a perfect detection algorithm and remove any bias from errors in the detection algorithm (k_d and k_r). Also, in the laboratory dataset there are no biases from the vine (k_v) or from other clusters (k_c) and hence we can isolate and study the bias from self-occlusions (k_b).

Initially we compare the total berry count (gathered manually) of each cluster against its weight, Table 2. The correlation score for total berry count to weight is $r^2 = 0.95$ and mean squared error from least squares fit of 9.3%. We consider this an upper bound for the yield predictions as the best yield prediction we could achieve is accurately knowing the berry count.

Table 2: Measurement Correlation to Fruit Weight

Measure-type	r^2 correlation	Mean squared error
Total berry count	0.95	9.3%
Visible berry count (Eq. 5)	0.88	15.4%
Ellipsoid model (Eq. 7)	0.85	17%
Convex hull model (Eq. 6)	0.92	13.7%

Next we study different image measurements starting with the visible berry count and present the results in Table 2. The visible berry count correlates with $r^2 = 0.88$ which provides a mean squared error of 15.4%. The error is just 6% greater than the total berry count and indicates a similar fraction of visible berries is present for small clusters as with large. The ellipsoidal model has a correlation score of $r^2 = 0.88$ and the lowest mean squared error of 17%. Even though the ellipsoidal model attempts to predict the occluded berries behind the visible layer of berries, it correlates with a lower score than the visible berry measure. The ellipsoidal model could be less accurate because it violates one of our assumptions; the clusters do not have uniform density or the clusters are not ellipsoidal, or the model could suffer from errors in the designation of the cluster contour.

The final image measurement model we evaluate is the convex hull in Tab. 2. The correlation measures at $r^2 = 0.92$ which is the best of the three image measurements we study. One possible reason for the high correlation is because it encompasses the entire cluster contour, therefore it includes a measure of the partially visible berries as well as the completely visible berries, thus being more accurate than visible berry count alone. Despite finding that the contour area in the image is more accurate measure other than visible berry count, we do not yet deploy this measure outside the laboratory environment. In the datasets collected in the vineyards, several clusters are visible in each image and we have yet to develop a technique for successful segmenting one cluster from another – a requirement of the ellipsoid and convex hull models. Hence, at present we have only been able to demonstrate precise detection of individual berries, regardless of which cluster they belong, and therefore in the following vineyard results we consider just the visible berry count that our algorithm provides. However, we have indicated with these laboratory tests an avenue for future improvement of our fielded system.

5.3 Biases in Visual Measurement

Here we take data both collected in the vineyard and laboratory we attempt to segregate the biases involved in the visual detection process. Table 3 presents a study of the different visual estimation biases measured in the various datasets. The laboratory dataset was used to estimate the self-occlusion parameter k_s by manually counting visible berries in the images and hand counting the total number of berries

physically pulling apart the clusters. The dataset showed that 46% of the berries are visible to the camera.

Table 3: Biases in Visual Yield Estimation (Sec. 3). Bias parameters are unit-less and standard deviations where applicable are presented as percentages in brackets. Missing entries for the self-occlusion parameter (k_s) are due to data not collected in a controlled environment in the vineyard datasets, the other missing entries are for the biases not applicable to the laboratory dataset.

Dataset name	Mean berry weight at harvest	k_s	k_d	k_r	$k_v * k_c$
Riesling	1.5g	-	0.74 (13.4%)	1.0	0.29
Traminette	1.6g	-	0.75 (10.7%)	1.0	0.24
Chardonnay	0.9g	-	0.65 (12.3%)	1.03 (8.7%)	0.31
Laboratory	-	0.46 (22%)	-	-	-

The visual detection bias k_d was estimated by manually assessing false positive and false negative detections of berries in images with the marked output of the detection algorithm. The false positive and false negatives combined with the true berry detections gave us the parameter k_d . The detection algorithm is biased towards under-counting the berries which was discovered in our prior work Nuske et. al [9] and here we see between 25% and 35% fewer berries than are visible are reported by the algorithm. The algorithm detected about 10% fewer of the visible berries in the Chardonnay 2011 dataset which captured around 90 days from harvest when the berries were much smaller. Similarly we take images marked by the detection and image alignment algorithm and manually assess double-counting and mis-counting berries between overlapping images. In the 2011 Chardonnay dataset the automatic aligning algorithm was biased to slightly over-count berries by 3%. In the 2010 datasets (Riesling and Traminette) we did not deploy our automatic alignment algorithm, and instead manually cropped images to have zero overlap and hence we list a bias of 1.0 here.

Finally, the bias from vine and cluster occlusions we have not experimentally determined, but we combine the other bias factors, assume a similar self-occlusion rate to the laboratory, use the mean berry weight at harvest and deduce the combined effect of the terms k_v and k_c indirectly. We see here that there are significant occlusions from the vine leaves, shoots and clusters occluding clusters with between two thirds and three quarters of the fruit being occluded by these factors.

5.4 Yield Prediction Results

Now we present yield prediction results by applying our calibration approaches described in Section 4, including calibration from destructive hand samples at the time

of imaging and also calibration from prior harvest data and compare average error for the prediction of the individual vine weights of the Chardonnay dataset. After discovering in the previous section that the visible berry count has the most accurate correlation to yield in the vineyard datasets we use this as our image measure for the following results. We take the Traminette and Riesling datasets, collected in 2010,

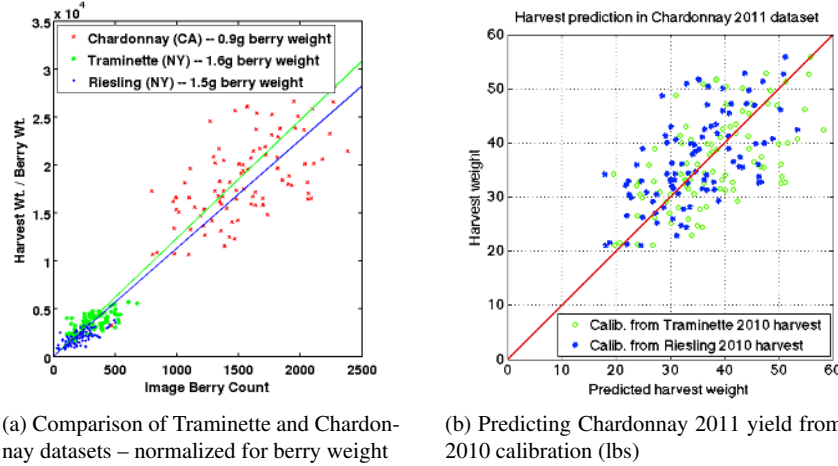


Fig. 3: Graphs demonstrating calibration from prior harvest data. After normalizing for berry weight there is a trend between the 2010 and 2011 datasets. The calibration functions computed from the 2010 datasets applied to the Chardonnay image berry counts produces harvest predictions 12 weeks prior to harvest.

compute a calibration between image measure and yield, and apply to the Chardonnay dataset from 2011. Fig. 3a shows a comparison between the data collected in the two vineyards after normalizing for respective berry weights. The graph shows that the Traminette and Riesling vines despite holding much less fruit, do have a trend between the image berry counts in the Chardonnay data. We apply the Traminette and Riesling calibration to the Chardonnay data and show the predicted weight in Fig. 3b. This result demonstrates the prediction of harvest yield 12 weeks out from harvest. We also evaluate calibrating using hand samples collected in the vineyard at the time of imaging, see Fig. 2b.

The results of the calibration on the Chardonnay dataset are presented in Fig. 4a. The errors are between 17 and 19% for the different calibration approaches, with the calibration from destructive hand samples slightly more accurate. We see that some of the error averages out when comparing yield of entire rows, where the error is now between 7% and 8%. In Fig. 4b we present the error for prediction of the entire yield of the vines in the dataset. The hand calibration was most accurate at 3% error and using calibration from the 2010 Traminette dataset had 4% error. We see an under-prediction of overall weight by 4.5% using the 2010 Riesling dataset

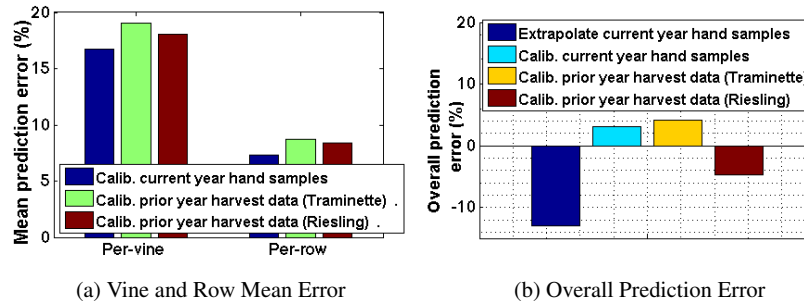


Fig. 4: Results on the prediction of harvest yield in the Chardonnay dataset. We evaluate the prediction accuracy when calibrating using a destructive hand sample at the time of imaging compared to calibrating using prior harvest data. We present two statistics in (a); the average absolute error computed for individual vine predictions and the average absolute error for the estimate of row weights. In (b) we present the error for prediction of the entire yield of the vines in the dataset. The calibration from hand samples is slightly more accurate than the calibration from prior harvest data. For comparison we present the estimate taken by extrapolating the hand samples alone, with is the least accurate estimate with 13% error.

calibration. It is apparent that despite average absolute per-vine errors of around 18% for all approaches, the overall error is below 5% indicating the individual vine errors are well distributed and average out. For comparison in Fig. 4b we present the estimate taken by extrapolating the hand samples alone, which is the traditional industry practice, and was found to be the least accurate estimate with -13% error.

6 Conclusion and Future Work

We have demonstrated a method to automatically generate non-destructive, high-resolution, yield predictions vineyards and shown practical approaches to calibrate the measurements. We evaluate the system at both 10 days from harvest and 90 days, finding similar results, with the algorithm detecting only slightly less of the visible berries earlier in the season. Of the two methods of calibrating image measurements – using destructive hand samples at the time of imaging was slightly more accurate at giving predictions within 3% of harvest yield. Nevertheless calibration from prior harvest data from a different varietal, albeit grown in a similar vine trellis and training structure, gave only 4% error.

In future work we will look to develop an approach to count grape clusters early in the season, even before berries have formed, to give vineyard managers information with maximum time before harvest to make the necessary adjustments to their

vines. We also look at extending the experimentation to more varieties and trellis structures.

7 Acknowledgements

Work funded by the National Grape and Wine Initiative, (info@NGWI.org). Narasimhan was supported partially by NSF awards IIS-0964562 and CAREER IIS-0643628 and an ONR grant N00014-11-1-0295.

References

1. Blom, P., Tarara, J.: Trellis tension monitoring improves yield estimation in vineyards. *HortScience* **44**, 678–685 (2009)
2. Clingeleffer, P., Dunn, G., Krstic, M., Martin, S.: Crop development, crop estimation and crop control to secure quality and production of major wine grape varieties: A national approach. Tech. rep., Grape and Wine Research and Development Corporation, Australia (2001)
3. Dey, D., Mummert, L., Sukthankar, R.: Classification of plant structures from uncalibrated image sequences. In: *IEEE Workshop on the Applications of Computer Vision (WACV)* (2012)
4. Dunn, G., Martin, S.: Yield prediction from digital image analysis: A technique with potential for vineyard assessments prior to harvest. *Australian Journal of Grape and Wine Research* **10**, 196–198 (2004)
5. Federici, J., Wample, R., Rodriguez, D., Mukherjee, S.: Application of terahertz gouy phase shift from curved surfaces for estimation of crop yield. In *Applied Optics* **48**, 1382–1388 (2009)
6. J.A., M.C., X., B.: Viticultura de precisión: Predicción de cosecha a partir de variables del cultivo e índices de vegetación. *Revista de Teledetección* **24**, 67–71 (2005)
7. Jimenez, A., Ceres, R., Pons, J.: A survey of computer vision methods for locating fruit on trees. In: *Transaction of the ASAE*, vol. 43, pp. 1911–1920 (2000)
8. Kitt, B., Geiger, A., Lategahn, H.: Visual odometry based on stereo image sequences with ransac-based outlier rejection scheme. In: *IEEE Intelligent Vehicles Symposium* (2010)
9. Nuske, S., Achar, S., Bates, T., Narasimhan, S., Singh, S.: Yield estimation in vineyards by visual grape detection. In: *Proceedings of the 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems* (2011)
10. Singh, S., Bergerman, M., Cannons, J., Grocholsky, B., Hamner, B., Holguin, G., Hull, L., Jones, V., Kantor, G., Koselka, H., Li, G., Owen, J., Park, J., Shi, W., Teza, J.: Comprehensive automation for specialty crops: Year 1 results and lessons learned. *Journal of Intelligent Service Robotics, Special Issue on Agricultural Robotics* **3**(4), 245–262 (2010)
11. Swanson, M., Dima, C., Stentz, A.: A multi-modal system for yield prediction in citrus trees. In: *ASABE Annual International Meeting*, Pittsburgh, PA (2010)
12. Wolpert, J.A., Vilas, E.P.: Estimating vineyard yields: Introduction to a simple, two-step method. *American Journal of Enology and Viticulture* **43**, 384–388 (1992)