2024 IEEE International Conference on Robotics and Automation (ICRA 2024)
May 13-17, 2024. Yokohama, Japan

Inexpensive, Automated Pruning Weight Estimation in Vineyards

Jonathan Jaramillo[1], Aaron Wilhelm[1], Nils Napp[1], Justine Vanden Heuvel[2], and Kirstin Petersen[1]

***Abstract*— Pruning weight is indicative of a vine's ability to produce a crop the following year, informing vineyard management. Current methods for estimating pruning weight are costly, laborious, and/or require specialized know-how and equipment. In this paper we demonstrate an affordable, simple, computer vision-based method to measure pruning weight using a smartphone camera and structured light which produces results better than state-of-the-art techniques for vertical shoot position (VSP) vines and demonstrate initial steps towards estimating pruning weight in high cordon procumbent (HC) vines such as Concord. The simplicity and affordability of this technique lends its self to deployment by farmers today or on future viticulture robotics platforms. We achieved an R2=.80 for VSP vines (better than state-of-the-art computer vision-based methods) and R2=.29 for HC vines (not previously attempted with computer vision-based methods).**

I. Introduction

The US wine and grape industry accounts for 1.28% of the US GDP ($276B) and contributes ∼1.84M US jobs with over one million acres of land dedicated to grape production [1]. Given their financial impact and the fact that grapes are the highest value crop produced in the US, there has been a push in recent years to adapt precision agriculture practices to the field of viticulture.

Precision viticulture is a management method designed to optimize yield, costs, and sustainability by accounting for inter- and intra-vineyard variables that are site and season specific. Deployment of precision viticulture practices rely on automated quantification of these variables using a variety of sensors and data collection techniques [2], [3]. Despite the significance of the grape industry as a whole to the US economy, the average vineyard size is less than 100 acres and many medium to small vineyards lack the financial resources needed to integrate precision viticulture into their management practices. For instance, high labor costs and labor shortages can make manual data collection impractical and unaffordable while emerging technologies such as LiDAR, drones, and multispectral sensing are often too expensive to justify deployment on a smaller scale. Other sensing techniques, such as remote sensing using satellite imagery, lack the spatial resolution needed to make localized management decisions on a vine by vine or block by block basis [4]. To this end, there is a need to make advancements in automated vineyard measurements for both farmer operated and robotic sensing platforms that are sufficiently spatially sensitive to inform decision making at this scale. Such measurements include yield estimation, normalized difference vegetation index (NDVI), leaf area index (LAI), crop coefficient, and pruning weight to name a few. These sensing methods need to be 1) affordable enough to justify deployment on a small to medium scale vineyard, 2) easy to use without the need for specialized training or technical expertise, 3) robust and reconfigurable enough for both use by farmers today and deployment on future robotics platforms that are becoming more ubiquitous in modern farming.

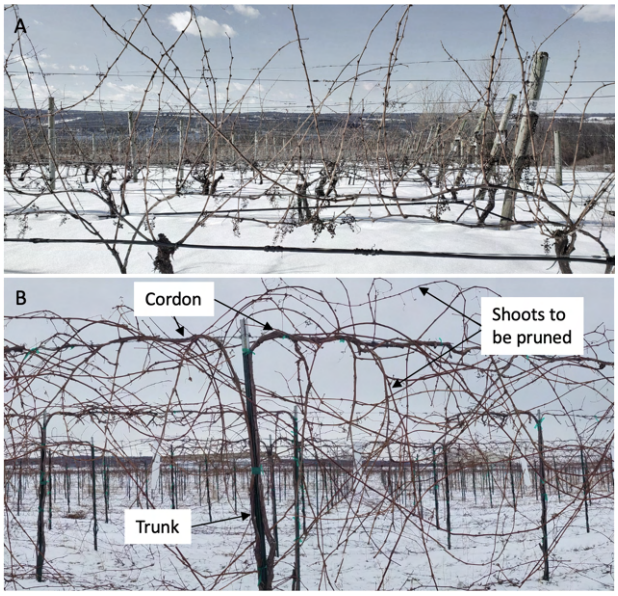

Fig. 1. A) VSP vines grow upwards starting from a cordon or cane that is ∼1m above the ground. B) HC Concord grapes sprawl in all directions off the cordon at a height of ∼1/.5m, many of the shoots grow downwards.

Pruning weight is the measurement of one year old growth removed after dormant pruning and is an effective way to assess biomass production, vine balance, vigor, and the carbon storage cycle in grapevines [5], [6], [7] (Fig. 1.A-B). Most importantly, the pruning weight is indicative of the vine's ability to produce a crop the following year. The ratio of yield to pruning weight is called the Ravaz Index [6], and for most vineyards has an optimal value of 5-10 [8]. In conjunction with the previous seasons' yield, each year growers can use pruning weight to calculate the Ravaz Index (Eq. 1) and adjust the number of buds per vine during cropping to optimize their vine balance for the next year.

$$Ravaz\ Index = \frac{Yield}{Pruning\ Weight} \tag{1}$$

Obtaining optimal vine balance is particularly important in colder climates with shorter growing seasons where over-cropping vines can result in under-ripened fruit. Moreover,

Funding for this research was provided by NSF grant #1837367 and a NIFA grant #1014705.

[1]Electrical and Computer Engineering, Cornell University, Ithaca, NY 14853, USA. [2]Plant Sciences, Cornell University, Ithaca, NY 12853, USA. jdj78@cornell.edu, ajw344@cornell.edu, nnapp@cornell.edu, justine@cornell.edu, kirstin@cornell.edu

pruning weights can be used in experimental research and to inform management practices beyond pruning, such as rootstock selection [9], nitrogen fertilization [10], phosphorus fertilization [11], and irrigation [12]. Manual methods for pruning weight data collection are laborious, time consuming, and, in the North East, can demand extended exposure to extreme cold weather during the dormant months of winter while interfering with the usual pruning workflow [13]. These factors often deter growers from collecting the large amounts of data needed to calculate a more accurate vineyard wide average Ravaz Index and make estimating the Ravaz Index on a vine-by-vine or panel-by-panel basis almost impossible. In many vineyards, these factors completely deter the use of Ravaz Index to inform management practices altogether. This illustrates the opportunity for automation in pruning weight estimation to lend itself to increased accessibility of precision viticulture practices to grape growers. In this paper we demonstrate an affordable, simple method to measure pruning weight using a smartphone camera which produces results on par with state-of-the-art techniques.

II. Related Work

Advances in computer vision (CV) have led to breakthroughs for autonomous crop monitoring systems on future robotics platforms. A variety of such robots are at or near commercial availability for vineyards, such as VineScout, GRAPE, and Burro. Numerous ground-based systems have been developed to estimate yield [14], [15], disease detection [16], [17], pruning and shoot characterization [18], cluster compactness [19], and seed inspection [20]. A more comprehensive list of ground-based CV systems in precision viticulture has been compiled by Seng et al. [21]. Similarly, drones equipped with multispectral camera sensors are capable of measuring the NDVI in vineyards [22], [23], and RGB videos from drones leverage photogrammetry to visualize and quantify vineyard canopy leaf area index [24] and to predict soil water erosion in Mediterranean vineyards [25].

In recent years, a variety of sensing methods have focused specifically on pruning weight estimation using both air and ground-based systems. Dobrowski et al. [26] used multispectral aerial images of 0.5m/pxl spatial resolution to estimate pruning weight density (kg/m) in sections of $\sim$15m with R2=0.68 and R2=0.88 over two years. Garcia-Fernandez et al. [27] achieved an R2=0.62 on blocks of two vines using 3D point clouds generated via photogrammetry from drone imagery. With the advancement of LiDAR sensors, a variety of ground-based 3D sampling platforms have been developed for estimating pruning weight. GRover [28] achieved a R2=0.91 for pruning weight and R2=0.76 for wood volume. While ground based LiDAR systems can provide a higher level of detail than aerial systems for measuring biomass, these systems are expensive; costing on the order of thousands of dollars for the LiDAR sensor alone [29].

Efforts to deploy lower cost sensors in the form of simpler RGB cameras started with McFarlane [30], although they concluded that more work was needed to determine pruning weight. Botterill [31], [32] developed an automatic pruning machine that travels down both sides of the vine at 2min/vine. It captures a 3D model of the vine, and mechanically prunes using a 6 joint robotic arm. The fundamental issue of image segmentation and uncontrolled lighting conditions in the field, however, is solved by completely enclosing the vine in a box. An alternative solution for image segmentation is to use a white sheet on the far side of the vine [33], [34], [18], [18], or stereoscopic imaging [33], [18]. Throughout this paper we will refer to ground based sensor systems that don't require backdrops or vine enclosure that can be mounted to a vehicle and driven through the vineyard as "on-the-go" systems. Millan et al. [34] achieved on-the-go RGB camera pruning weight estimations without the use of a backdrop and R2=0.77. Segmentation was achieved using a supervised Mahalanobis distance classifier and a support vector machine for image segmentation. This methodology requires a $1000 mirrorless DSLR camera and must be deployed on an ATV with custom mounting and electric odometric triggering hardware which requires specialized training. This system is highly sensitive to lighting parameters. These factors can be prohibitive to grape growers looking to deploy automated pruning weight measurements both in terms of cost and know-how.

The majority of vision-based automated pruning weight assessment have been focused on vertical shoot positioned vines, with little attention given to high cordon (HC) vines. Researchers working towards automation of pruning weight for HC vines have exclusively used aerial systems or relied on more expensive LiDAR sensors, as opposed to cheaper RGB cameras. Vertical shoot positioned (VSP) vines have a more planar, upward growing structure and can more accurately be represented with a 2D model, while HC vines are less structured with shoots pointing in all directions and draping towards the ground, which explains why 3D sensor and vine models are needed to measure pruning weight.

We propose a different approach to CV pruning weight estimation that uses structured light and a smartphone camera in an effort to address issues of affordability and ease of use, lowering the barrier of entry for farmers and horticulturalists. Moreover, simplification of hardware and increasing robustness to lighting parameters by eliminating light sensitive machine learning models [35] will lend itself to deployment on fully autonomous vineyard monitoring robotic systems. In previous work [36] we demonstrated the potential for smartphone cameras to be used in vineyards for yield estimation. One advantage of this approach is the affordability and ubiquitousness of smartphones. Many grape growers already own a smartphone, eliminating the need to purchase expensive equipment as well as the know-how needed to operate this equipment. The aim of this work is to develop a robust smartphone-based CV pipeline using simple structured light to assess pruning weights for use by farmers today, and to lay the groundwork for deployment on future autonomous robots. The sensor system is designed specifically to be inexpensive, simple to use, and compact to be easily deployed by manually walking down a row, attaching it to a tractor or ATV, or on a robotics platform

such as TerraSentia (Earth Sense) or GRover [28]. We have developed a CV pipeline capable of outperforming state of the art on-the-go CV-based pruning weight estimation for VSP vines, and demonstrate first steps towards an adaptation of this method to be used for HC vines.

III. Data Collection

Data for VSP vines was collected at the Cornell teaching vineyard in Lansing, New York (42° 34' 22.32"N, 76° 35'48.22"W). Mature Vitis vinifera L. (Riesling and Cabernet Franc) with vine spacing of 6ft x 9ft were cane pruned and vertically shoot-positioned in accordance with local practices [5]. Video data was collected in mid-March, and subsequent pruning weight data was collected at the end of March. Data for HC canopies were collected at the Cornell Lake Erie Cooperative Extension Research Facility (42°22' 19.4592"N, 79°29' 8.3538"W). Mature Vitis labrusca (Concord) with vine spacing of 8ft x 9ft were spur pruned. To serve as ground truth data, pruning weights were manually collected: the vines were pruned, the clippings were bundled, tied, and weighed with a small digital scale. Video data was collected at the beginning of December and subsequent pruning weight data collection was completed by the end of December, a time of the year where temperatures average below freezing in many cool climate regions like New York.

The system consists of a phone, a small battery, and a line laser. The green line laser (VLM-520-28 LPT) cost $28 and is a Class 1 laser, producing less than 0.39mW of light, with ∼510-530nm wavelength. This laser was selected because of its relatively low cost, safety (no need for eye protective equipment), and emitting angle of >60°. The laser is powered with a 9V battery and is mounted to the phone via a 3D printed carrying case shown in Fig. 2.A. The laser is mounted 0.23m to the left and 0.13m in front of the iPhone camera at a 30° angle relative to the field of view. This horizontal offset lends itself to determining which row of vines the laser is hitting (Fig. 2.B). The entire system is mounted on a stabilizing gimbal (Zhiyun-Tech Weebill-S). An LED light (Neewer CN-160) covered with a red piece of cellophane was also mounted to the handle of the gimbal to create red ambient light for easy detection/differentiation of the green laser light in the video processing pipeline.

Videos were collected at night using an iPhone 13 Pro with a frame rate set to 240fps (Fig. 2.C-E). In actuality, the phone collected videos at ∼180fps, likely due to low lighting forcing higher exposure times. The gimbal was carried by hand down the row at a brisk walking pace (roughly 1.25 min/row) with the system facing perpendicular to the vine canopy. This pace and frame rate corresponds to a reconstruction spatial resolution of ∼5mm. It could easily be mounted to an ATV or tractor depending on preference. The resulting videos were downloaded from the cloud to be processed on a desktop computer.

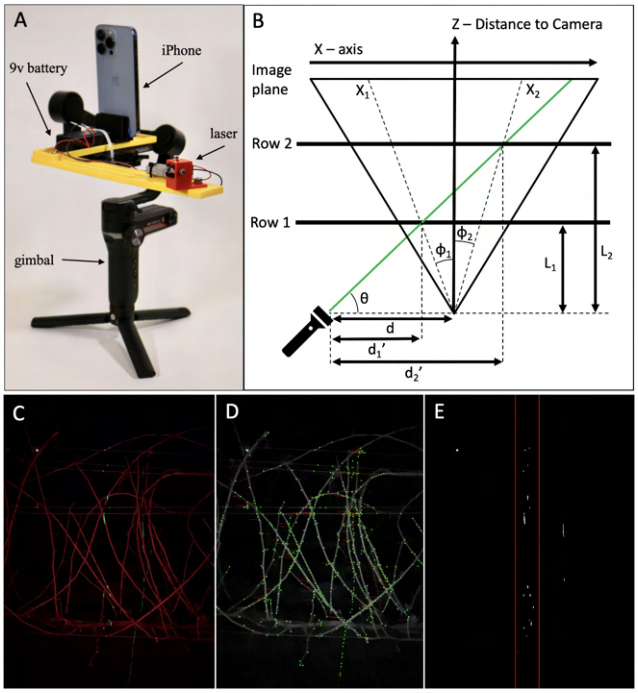

Fig. 2. A) The iPhone is mounted to a gimbal via 3D printed L-shaped part. All non-3D printed hardware is off-the-shelf and simple to assemble. B) The horizontal distance between the camera and the laser (d), and the angle of the laser relative to the camera (θ) can be used to determine the distance of a pixel from the camera given the horizontal coordinate of the pixel using a camera pinhole model. C) Example of video captured by iPhone camera at night. D) Image features are detected and tracked in the red color channel for estimating camera motion. E) A threshold is applied to the green color channel and only reflected light in the region of interest corresponding to the nearest row is used in the vine reconstruction.

IV. Data Processing

A. VSP Vine Pipeline

The CV pipeline generates a 2D silhouette of the vine and subsequently determines the pruning weight by counting the pixels corresponding to the vine. Green structured laser light and red ambient light were selected to simplify the data processing steps. Isolating the structured light from the unstructured light is trivially achieved by isolating the green color channel. The video processing pipeline is composed of three unique stages. 1) Preprocessing, 2) Tracking and reconstruction, 3) Image analysis (Fig. 3.A).

Stage one determines the region of the video that contains patterns of light reflected from the row of vines closest to the camera. The position of the laser is offset to the left and panned right relative to the camera. From the camera frame of reference, objects closer to the camera will create light patterns further left on the camera's projection screen and objects further from the camera will create patterns further right (Fig. 2.B). Rows of VSP vines are modeled as planes that are discrete distances from the camera, therefore, light patterns will be clustered along the horizontal axis depending on which row the light is reflecting from. The angle of the laser relative to the camera is set such that objects further than two row spacings from the camera will not be illuminated within the camera's field of view.

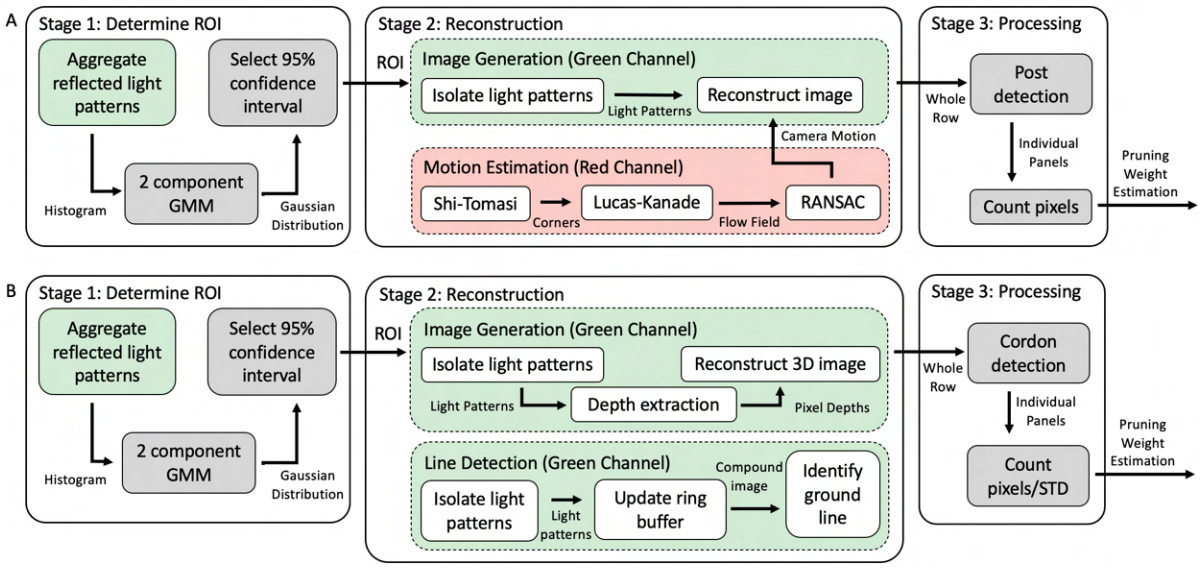

Fig. 3. A) Overview of the CV pipeline for VSP vines. B) Overview of the CV pipeline for HC vines.

To select the horizontal region corresponding to the nearest row, 2,000 frames in the middle of the video are selected, the green channel of the images are isolated, and the laser light pattern is detected using thresholding. The horizontal location of all laser reflections across all frames is recorded and a histogram is generated (Fig. 4.A). As expected, this histogram has a bimodal distribution, the leftmost peak corresponding to the closest row, and the rightmost peak corresponding to the further row. A Gaussian mixture model with two components is fitted to the histogram, the mean and standard deviation of each distribution is calculated, and a horizontal window corresponding to a 95% confidence interval for the first row is selected. This confidence interval corresponds to the area between the red lines shown in Fig. 2.E. Any light patterns outside of this region are ignored as they correspond to the row of vines in the background.

The second stage is camera tracking and image reconstruction. Tracking the camera motion requires isolating the red channel (Fig 2.C) and OpenCV's implementation of the Shi-Tomasi corner detection algorithm (a variation on Harris corner detection) [37] is used to find a set of features for each frame as shown in Fig 2.D. These features are tracked from frame to frame using OpenCV's implementation of Lucas-Kanade optical flow tracking [38]. A 2D transformation is calculated from frame to frame using this optical flow field. Outliers, such as features detected in the background, or as a result of artifacts from the green laser bleeding into the red channel, are removed using RANSAC [39] and a 2D translation in camera position from frame to frame is calculated. A 2D translational motion model yields acceptable results as the gimbal removes most rotation, and distance to the vine is relatively consistent throughout the video.

Reconstructing an image of the vine is achieved by isolating the light patterns from the green channel that are within the region of interest and collapsing them into a single binary

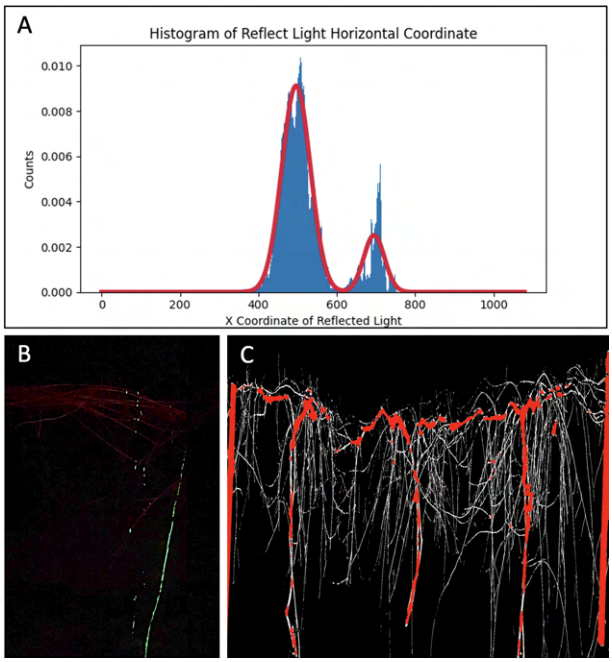

Fig. 4. A) Histogram (blue) of the horizontal position of every white pixel in the thresholded green channel for a uniformly selected subset of video frames. A two-component Gaussian mixture model (red) is fitted to determine the region of interest. B) For HC vines the camera is pointed downwards and a line is visible as a result of the ground being in the camera frame. C) Post and cordon/trunk detection in Concord vines.

column. One or zero corresponds to the presence or absence of light patterns at a vertical position using a binary OR operator. The column is then shifted vertically and scaled horizontally based on the translation of the optical flow field generated from the tracking/motion estimation. The width of the column is determined by the horizontal displacement of the optical flow field. For this reason tracking and reconstruction are performed in tandem. The frame of the video

generates a new column and these columns are horizontally concatenated to generate a silhouette reconstruction of the vine (Fig. 5).

The last stage segments the silhouette image of the entire row into individual panels along post boundaries. First, posts are detected using a spatial low pass filter and thresholding. The image is segmented along the major axes of each post's pixel region and the pixels corresponding to the post are removed. Because the cordon is removed each year and considered part of the pruning weight, no cordon detection is implemented. Similarly, the trellis wires are included as part of the reconstruction and are not removed from the image because every panel contains the same amount of wire. The wire surface area can be accounted for in the y-intercept of the calibration line. The total number pixels in each image are counted and their sum is calibrated to estimate the pruning weight of each panel.

B. HC Vine Pipeline

The data processing pipeline for videos of HC vines is similar to that of the VSP vines with a few differences (Fig. 3.B). Unlike VSP vines that grow upward, the camera had to be pointed downwards to capture the lower hanging shoots, which makes the ground visible in the frame and the laser reflection off the ground creates a straight line pattern. This ground line (Fig. 4.B) is detected, such that it can be ignored, using the following steps.

To isolate the ground line, a ring buffer of 5 frames is created. For each of the frames in the buffer, the green channel is isolated and a threshold is applied to create a binary image. Then all the binary images in the ring buffer were aggregated into a single image using a pixel-wise logical OR operator. A linear Hough transform is applied to find the dominant line created by the laser on the ground. Because this method is noisy, occasionally dropping frames or generating false lines, a temporal filter is applied that removes outliers by assuming that the ground line will change very slowly from frame to frame. Any frame whose line lies too far from the location of the previous five lines is considered to be an outlier and is replaced by the coordinates or the line determined in the previous frame. All light reflected to the right of this line is ignored. Determining the vertical bounds between rows is then completed using the same steps as the VSP vines.

Because HC vines have a less planar structure than VSP vines, the model incorporates some of the 3D structure of the vine. In the second stage the mean horizontal pixel location is stored for each pixel of the resulting column. A camera model (Fig. 2.B) given by Eq. 2 can be used to convert these horizontal pixel locations to their corresponding distance from the camera. Concatenating these columns results in an image that is a 3D reconstruction of the vine, with each pixel value corresponding to the distance of the vine from the camera.

$$L = \frac{d}{\frac{1}{\tan\theta} - \frac{x \tan HFOV}{w}} \tag{2}$$

Where L is the distance of the pixel/vine from the camera, θ is the angle of the laser relative to the camera, d is the horizontal offset between the laser and the camera, $HFOV$ is the horizontal field of view of the camera in degrees, w is the width of the image in pixels, and x is the horizontal pixel location of the reflected light.

The image analysis step detects posts and segmenting the image into individual panels. The cordons and trunks are detected using a similar technique to post detection as shown in Fig. 4.C. Each panel image is then partitioned into 400 smaller equally sized subimages, a value that was determined empirically given the resolution of the camera and walking speed. For each subimage, the number of pixels is counted and multiplied by the standard deviation of the pixel values in that subimage. These values are summed to generate the pruning weight estimation. Our reasoning for this computation is that regions of the image that have more pixels contain more pruning weight mass. Likewise, regions of the image that have a higher spread in their distance to the camera contain more pruning weight due to the shoots being more perpendicular to the camera plane.

V. RESULTS

TABLE I: Comparison of pruning weight estimation methods. Bolded entries are results from this paper. OTG: on-the-go. TS: training system.

	Sensing Method	Cost	Platform	OTG	Resolution (vines)	Dataset Size	TS	Results (R2)
[26]	Multispectral Imaging	$25k-$100k [40]	Plane	NA	∼0.5	∼90	VSP	0.88
[27]	Photogrammetry	$1600	UAV	NA	2	20	?	0.62
[28]	LiDAR	$4500	Custom Rover	Yes	3	11	Both	0.91
[33]	Stereoscopic Image Segmentation	$8000	Custom Rover	Yes	1	39	VSP	0.23
[33]	Manual Image Segmentation	$8000	Custom Rover	Yes	1	39	VSP	0.84
[34]	Backdrop Image Segmentation	$2000	Modified ATV	No	1	44	VSP	0.92
[34]	Mahalanobis distance classifier & SVM	$2000	Modified ATV	Yes	1	44	VSP	0.77
*	**Structured Light 2D**	**$500**	**ATV/ Walk**	**Yes**	**3**	**55**	**VSP**	**0.80**
*	**Structured Light 3D**	**$500**	**ATV/ Walk**	**Yes**	**4**	**216**	**HC**	**0.80**

A. Statistical Analysis

The pruning weight estimates are fitted to measured weights using least squares linear regression. Leave one out cross validation (LOOCV) is a common method used to evaluate the performance of an estimation model on smaller datasets. LOOCV is performed by removing one data point from the data set, fitting a regression line, and using that regression line to predict the pruning weight of the excluded data point. This process is repeated for each data point in the data set. We use this technique to calculate a residual sum of squares (R2), root mean squared error (RMSE), and mean absolute error (MAE).

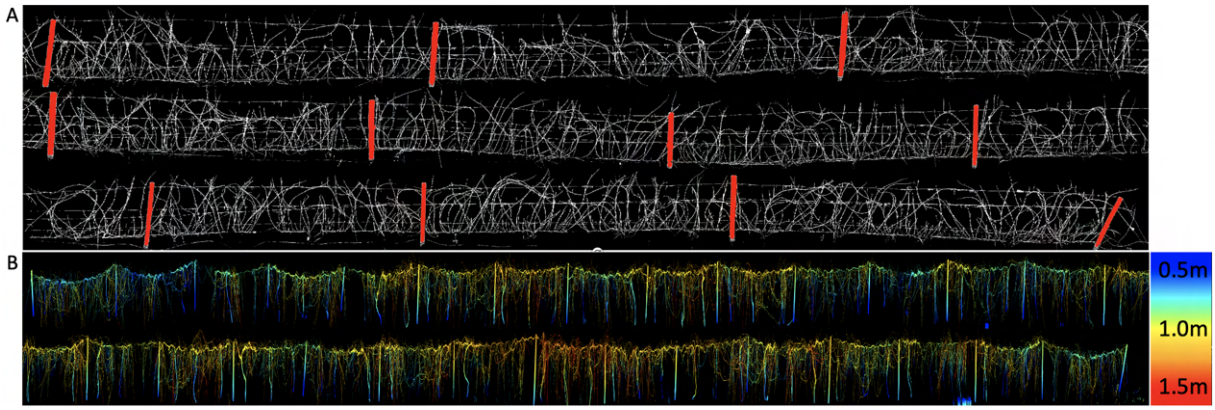

Fig. 5. A) Reconstruction of VSP vines; the planar structure of VSP vines lends itself to modeling with a 2D structure. B) Reconstruction of the HC vines with 3D model. The color scale below shows the estimated distance from the camera for each pixel.

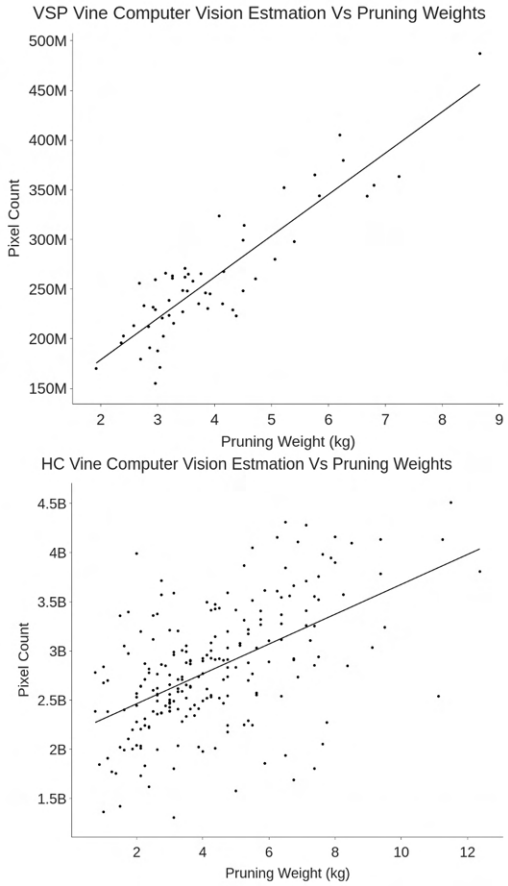

Fig. 6. A) 55 panels of VSP vines pruning weight estimations are plotted against manually measured pruning weights. Linear regression line: $y = 4.16e7x + 9.56e7$ (R2=.80). B) 216 panels of HC vines pruning weight estimations are plotted against their manual measurements. Linear regression line: $y = 1.22e7x + 1.86e8$ (R2=.29).

B. VSP Vines

The VSP vine reconstruction is shown in Fig. 5.A with the CV estimations plotted against manual measurements in Fig. 6.A. The linear regression plot yielded R2=.80 and the LOOCV of the linear model yielded R2=.78, RMSE=290.9g, and MAE=245.3g.

C. HC Vines

The HC vine reconstruction is shown in Fig. 5.B along with the automated post detection results. The resulting pruning weight estimations are plotted against the measured weights in Fig. 6.B. The regression plot yielded R2=.29 and the leave one out cross validation of the linear model yielded R2=.27, RMSE=859.6g, and MAE=649.8g. We found measuring the depth variance of the 3D model to yield more accurate results than simply counting the pixels as is done for the VSP vines. Using a 2D reconstruction for HC pruning weight estimations (similar to the VSP techniques), our system achieved R2=.23.

VI. Discussion

We demonstrated that structured light outperforms current CV pruning weight estimation methods for VSP vines, and present easy and low-cost in-field deployment for farmers and future robots. To our knowledge this paper is the first attempt to use ground-based CV for estimating pruning weight of vines with a procumbent canopy. While our approach to HC canopies is not yet sufficient for immediate commercial applications, we have demonstrated that incorporating 3D information into the vine model improves pruning weight performance (R2=.30 with 3D model vs. R2=.23 with 2D model). Smartphones, unlike LiDAR systems, and drones, are ubiquitous. Moreover, our method does not rely on machine learning models that can be highly sensitive to light fluctuations and require large amounts of training data and labeling. Farmers can deploy this method without buying or learning to operate highly specialized equipment; they can likely use the smartphone that is already in their pocket. Moreover, smartphone cameras are relatively cheap and small, making them accessible for integration into small ground-based robots.

Acknowledgements

We would like to thank Terry Bates, Senior Research Associate at the Cornell School of Integrative Plant Science Horticulture Section and Director at the Cornell Lake Erie Research and Extension Lab Cornell AgriTech for providing pruning weight data for this paper.

REFERENCES

[1] J. D. . Associates, "2022 economic impact study of the american wine industry methodology and documentation," 2022.

[2] L. Mohimont, F. Alin, M. Rondeau, N. Gaveau, and L. A. Steffenel, "Computer vision and deep learning for precision viticulture," *Agronomy*, vol. 12, no. 10, p. 2463, 2022.

[3] J. Arnó Satorra, J. A. Martínez Casasnovas, M. Ribes Dasi, and J. R. Rosell Polo, "Precision viticulture. research topics, challenges and opportunities in site-specific vineyard management," *Spanish Journal of Agricultural Research, 2009, vol. 7, núm. 4, p. 779-790*, 2009.

[4] M. Roznik, M. Boyd, and L. Porth, "Improving crop yield estimation by applying higher resolution satellite ndvi imagery and high-resolution cropland masks," *Remote Sensing Applications: Society and Environment*, vol. 25, p. 100693, 2022.

[5] T. Wolf, "Wine grape production guide for eastern north america (nraes 145)," 2008.

[6] R. Smart, M. Robinson, *et al.*, *Sunlight into wine: a handbook for winegrape canopy management.* Winetitles, 1991.

[7] M. Keller, *The science of grapevines*. Academic press, 2020.

[8] B. Komm and M. Moyer, "Vineyard yield estimation," 2015.

[9] J. Lambert, M. Anderson, J. Wolpert, *et al.*, "Vineyard nutrient needs vary with rootstocks and soils," *California agriculture*, vol. 62, no. 4, pp. 202–207, 2008.

[10] S. Spayd, R. Wample, R. Stevens, R. Evans, and A. Kawakami, "Nitrogen fertilization of white riesling in washington: effects on petiole nutrient concentration, yield, yield components, and vegetative growth," *American Journal of Enology and Viticulture*, vol. 44, no. 4, pp. 378–386, 1993.

[11] P. Skinner, J. Cook, and M. Matthews, "Responses of grapevine cvs chenin blanc and chardonnay to phosphorus fertilizer applications under phosphorus-limited soil conditions," *Vitis*, vol. 27, no. 95-109, 1988.

[12] J. Morris and D. Cawthon, "Effect of irrigation, fruit load, and potassium fertilization on yield, quality, and petiole analysis of concord (vitis labrusca l.) grapes," *American Journal of Enology and Viticulture*, vol. 33, no. 3, pp. 145–148, 1982.

[13] J. A. Taylor and T. R. Bates, "Sampling and estimating average pruning weights in concord grapes," *American Journal of Enology and Viticulture*, vol. 63, no. 4, pp. 559–563, 2012.

[14] S. Nuske, S. Achar, T. Bates, S. Narasimhan, and S. Singh, "Yield estimation in vineyards by visual grape detection," in *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2011, pp. 2352–2358.

[15] S. Liu, S. Cossell, J. Tang, G. Dunn, and M. Whitty, "A computer vision system for early stage grape yield estimation based on shoot detection," *Computers and Electronics in Agriculture*, vol. 137, pp. 88–101, 2017.

[16] G. Li, Z. Ma, and H. Wang, "Image recognition of grape downy mildew and grape powdery mildew based on support vector machine," in *Computer and Computing Technologies in Agriculture V: 5th IFIP TC 5/SIG 5.1 Conference, CCTA 2011, Beijing, China, October 29-31, 2011, Proceedings, Part III 5*. Springer, 2012, pp. 151–162.

[17] S. S. Sannakki, V. S. Rajpurohit, V. Nargund, and P. Kulkarni, "Diagnosis and classification of grape leaf diseases using neural networks," in *2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT)*. IEEE, 2013, pp. 1–5.

[18] M. Gao, "Image processing and analysis for autonomous grapevine pruning." Ph.D. dissertation, The University of Aidelaide, 2011.

[19] S. Cubero, M.-P. Diago, J. Blasco, J. Tardáguila, J. M. Prats-Montalbán, J. Ibanez, J. Tello, and N. Aleixos, "A new method for assessment of bunch compactness using automated image analysis," *Australian Journal of Grape and Wine Research*, vol. 21, no. 1, pp. 101–109, 2015.

[20] F. Avila, M. Mora, and C. Fredes, "A method to estimate grape phenolic maturity based on seed images," *Computers and Electronics in Agriculture*, vol. 101, pp. 76–83, 2014.

[21] K. P. Seng, L.-M. Ang, L. M. Schmidtke, and S. Y. Rogiers, "Computer vision and machine learning for viticulture technology," *IEEE Access*, vol. 6, pp. 67 494–67 510, 2018.

[22] N. S. Goel, "Models of vegetation canopy reflectance and their use in estimation of biophysical parameters from reflectance data," *Remote sensing reviews*, vol. 4, no. 1, pp. 1–212, 1988.

[23] P. J. Zarco-Tejada, J. R. Miller, T. L. Noland, G. H. Mohammed, and P. H. Sampson, "Scaling-up and model inversion methods with narrowband optical indices for chlorophyll content estimation in closed forest canopies with hyperspectral data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, no. 7, pp. 1491–1507, 2001.

[24] A. J. Mathews and J. L. Jensen, "Visualizing and quantifying vineyard canopy lai using an unmanned aerial vehicle (uav) collected high density structure from motion point cloud," *Remote sensing*, vol. 5, no. 5, pp. 2164–2183, 2013.

[25] M. Prosdocimi, M. Burguet, S. Di Prima, G. Sofia, E. Terol, J. R. Comino, A. Cerdà, and P. Tarolli, "Rainfall simulation and structure-from-motion photogrammetry for the analysis of soil water erosion in mediterranean vineyards," *Science of the Total Environment*, vol. 574, pp. 204–215, 2017.

[26] S. Dobrowski, S. Ustin, and J. Wolpert, "Grapevine dormant pruning weight prediction using remotely sensed data," *Australian Journal of Grape and Wine Research*, vol. 9, no. 3, pp. 177–182, 2003.

[27] M. García-Fernández, E. Sanz-Ablanedo, D. Pereira-Obaya, and J. R. Rodríguez-Pérez, "Vineyard pruning weight prediction using 3d point clouds generated from uav imagery and structure from motion photogrammetry," *Agronomy*, vol. 11, no. 12, p. 2489, 2021.

[28] M. H. Siebers, E. J. Edwards, J. A. Jimenez-Berni, M. R. Thomas, M. Salim, and R. R. Walker, "Fast phenomics in vineyards: development of grover, the grapevine rover, and lidar for assessing grapevine traits in the field," *Sensors*, vol. 18, no. 9, p. 2924, 2018.

[29] D. Andújar, H. Moreno, J. M. Bengochea-Guevara, A. de Castro, and A. Ribeiro, "Aerial imagery or on-ground detection? an economic analysis for vineyard crops," *Computers and electronics in agriculture*, vol. 157, pp. 351–358, 2019.

[30] N. McFarlane, B. Tisseyre, C. Sinfort, R. Tillett, and F. Sevila, "Image analysis for pruning of long wood grape vines," *Journal of agricultural engineering research*, vol. 66, no. 2, pp. 111–119, 1997.

[31] T. Botterill, R. Green, and S. Mills, "Finding a vine's structure by bottom-up parsing of cane edges," in *2013 28th International Conference on Image and Vision Computing New Zealand (IVCNZ 2013)*. IEEE, 2013, pp. 112–117.

[32] T. Botterill, S. Paulin, R. Green, S. Williams, J. Lin, V. Saxton, S. Mills, X. Chen, and S. Corbett-Davies, "A robot system for pruning grape vines," *Journal of Field Robotics*, vol. 34, no. 6, pp. 1100–1122, 2017.

[33] A. Kicherer, M. Klodt, S. Sharifzadeh, D. Cremers, R. Töpfer, and K. Herzog, "Automatic image-based determination of pruning mass as a determinant for yield potential in grapevine management and breeding," *Australian Journal of Grape and Wine Research*, vol. 23, no. 1, pp. 120–124, 2017.

[34] B. Millan, M. P. Diago, A. Aquino, F. Palacios, and J. Tardaguila, "Vineyard pruning weight assessment by machine vision: towards an on-the-go measurement system: This article is published in cooperation with the 21th giesco international meeting, june 23-28 2019, thessaloniki, greece. guests editors: Stefanos koundouras and laurent torregrosa," *Oeno One*, vol. 53, no. 2, 2019.

[35] D. I. Patrício and R. Rieder, "Computer vision and artificial intelligence in precision agriculture for grain crops: A systematic review," *Computers and electronics in agriculture*, vol. 153, pp. 69–81, 2018.

[36] J. Jaramillo, J. Vanden Heuvel, and K. H. Petersen, "Low-cost, computer vision-based, prebloom cluster count prediction in vineyards," *Frontiers in Agronomy*, vol. 3, p. 8, 2021.

[37] C. Harris, M. Stephens, *et al.*, "A combined corner and edge detector," in *Alvey vision conference*, vol. 15, no. 50. Citeseer, 1988, pp. 10–5244.

[38] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in *IJCAI'81: 7th international joint conference on Artificial intelligence*, vol. 2, 1981, pp. 674–679.

[39] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, 1981.

[40] S.-H. Baek, I. Kim, D. Gutierrez, and M. H. Kim, "Compact single-shot hyperspectral imaging using a prism," *ACM Transactions on Graphics (TOG)*, vol. 36, no. 6, pp. 1–12, 2017.