

# A Hybrid Cable-Driven Robot for Non-Destructive Leafy Plant Monitoring and Mass Estimation using Structure from Motion

Gerry Chen<sup>1</sup>, Harsh Muriki<sup>1</sup>, Cédric Pradalier<sup>2</sup>, Yongsheng Chen<sup>3</sup>, and Frank Dellaert<sup>1</sup>

**Abstract**—We propose a novel hybrid cable-based robot with manipulator and camera for high-accuracy, medium-throughput plant monitoring in a vertical hydroponic farm and, as an example application, demonstrate non-destructive plant mass estimation. Plant monitoring with high temporal and spatial resolution is important to both farmers and researchers to detect anomalies and develop predictive models for plant growth. The availability of high-quality, off-the-shelf structure-from-motion (SfM) and photogrammetry packages has enabled a vibrant community of roboticists to apply computer vision for non-destructive plant monitoring. While existing approaches tend to focus on either high-throughput (e.g. satellite, unmanned aerial vehicle (UAV), vehicle-mounted, conveyor-belt imagery) or high-accuracy/robustness to occlusions (e.g. turn-table scanner or robot arm), we propose a middle-ground that achieves high accuracy with a medium-throughput, highly automated robot. Our design pairs the workspace scalability of a cable-driven parallel robot (CDPR) with the dexterity of a 4 degree-of-freedom (DoF) robot arm to autonomously image many plants from a variety of viewpoints. We describe our robot design and demonstrate it experimentally by collecting daily photographs of 54 plants from 64 viewpoints each. We show that our approach can produce scientifically useful measurements, operate fully autonomously after initial calibration, and produce better reconstructions and plant property estimates than those of over-canopy methods (e.g. UAV). As example applications, we show that our system can successfully estimate plant mass with a Mean Absolute Error (MAE) of 0.586g and, when used to perform hypothesis testing on the relationship between mass and age, produces p-values comparable to ground-truth data ( $p=0.0020$  and  $p=0.0016$ , respectively).

## I. INTRODUCTION

Non-destructive methods for estimating plant properties are of interest to both researchers and farmers to maximize crop yields and minimize resource utilization by developing more accurate growth models and monitoring plant health more precisely [1]. While traditional methods of measuring plant properties require harvesting the plant (e.g. placing on a scale or performing a nutrient analysis) [2], harvesting the entire plant makes it impossible to measure the same plant multiple times. Instead, many “replicates” must be planted to (1) harvest periodically and (2) obtain a larger sample size

per harvest to compensate for plant variation, both of which make developing plant models labor and resource inefficient. Non-destructive methods, which are defined by not requiring harvesting, for estimating plant mass (and other properties) with both high throughput and accuracy would enable more accurate plant modelling [3], [4].

Significant work has applied computer vision and/or robotics to non-destructive plant analysis and can be broadly categorized into 2D and 3D approaches [5]. Works focusing on aggregate field biomass rather than individual plant mass [6] do not achieve comparable accuracies to those dedicated to individual plants, so we focus on the latter. 2D computer vision approaches with single-plant resolution monitor plant disease, health, and canopy area well [7], but struggle to estimate mass with high accuracy [8]: a critical component of growth modeling. 3D approaches include using RGB [9], Stereo [10], Time-of-Flight-based [11], [12], and a number of other specialized sensors to estimate plant properties such as mass, dimensions [13], and organ structure [14], [15]. In most cases, multiple viewpoints are required to accurately assess the 3D structure of a plant, with occlusions posing particular difficulty [13], [14]. Existing approaches tend to focus on either high-throughput or high-accuracy. High-throughput approaches, such as satellite [16], [17], UAV [14], [18]–[20], conveyor [21], and cart [22]–[26]-based systems, can phenotype large numbers of plants but are less capable of handling occlusions due to their limited viewing angles. Meanwhile, high-accuracy approaches collect data from many viewpoints using e.g. robot arms/gantries [27], [28], scanner turntables [29], and handheld scanners/photogrammetry [10], [11], but are not designed to image large numbers of plants in an automated fashion.

In this work, we propose a medium-throughput, high-accuracy approach that pairs the large workspace of a cable robot with the dexterity of a robot arm manipulator to collect photos of many plants from many viewpoints. We validated the hybrid cable-manipulator approach with a  $3\text{m} \times 2\text{m}$  cable-driven parallel robot (CDPR) and 4 degree-of-freedom (DoF) manipulator moving an RGB camera to image lettuce plants in the vertical hydroponic system shown in Fig. 1. We collected a dataset consisting of photos, fresh masses, and dehydrated masses of 54 plants. Finally, we demonstrated the scientific utility of the data collected by our robot both qualitatively and quantitatively by comparing dense 3D reconstructions, mass estimates, hypothesis tests, and occlusions against 3 baselines based on simulated photos for competing approaches.

This material partially supported by the National Science Foundation (Award No. 2008302), U.S. Department of Agriculture (Award No. 2018-68011-28371), and National Science Foundation–U.S. Department of Agriculture Award No. 2020-67021-31526.

<sup>1</sup>Institute for Robotics and Intelligent Machines, Georgia Tech, Atlanta, USA {gchen328, vmuriki3, fd27}@gatech.edu

<sup>2</sup>CNRS IRL 2958, Georgia Tech Lorraine, France cedric.pradalier@georgiatech-metz.fr

<sup>3</sup>School of Civil and Environmental Engineering, Georgia Tech, Atlanta, USA yongsheng.chen@ce.gatech.edu



Fig. 1. The robot is placed in a vertical indoor grow chamber between the plants and the grow lights (Left: side view; Right: top view). The robot consists of a CDPR which moves the robot arm (circled in red) from plant to plant to take photos from a wide set of viewpoints.

## II. APPROACH

The robot used for collecting photographic data of the lettuce plants consists of 2 subsystems: a robot arm mounted on the end effector of a cable-driven parallel robot (CDPR). The purpose of the robot arm is to collect large numbers of photos of a plant from various, repeatable angles for use in SfM and other analysis techniques. Meanwhile, the CDPR enables analyzing a larger quantity of plants by moving the robot arm from plant to plant, enlarging the workspace of the robot arm to cover dozens of plants.

### A. Mechanical Design

1) *CDPR*: The cable-based robot platform is chosen for its scalability and economy [30], [31], which allow the robot to reach multiple plants and remain permanently installed for complete autonomy. Although our demonstrated robot is only  $2.9\text{m} \times 2.3\text{m}$  in size, in principle it can scale to almost any size vertical grow towers. As compared to e.g. gantry or conveyor type systems, cable-based systems remain almost constant in price relative to size. The planar design is chosen for its favorable tradeoff between capability and cost/complexity, since collisions with plants would limit the utility of out-of-plane motions anyway.

The CDPR is an 8-cable, 4-motor planar CDPR with a workspace of roughly  $2.9\text{m} \times 2.3\text{m}$ . Details on the design can be found in [32], with the primary distinctions being that (a) the robot arm shown in Fig. 3 is used in place of the spray paint carriage, and (b) the cables are doubled to provide more out-of-plane stability. The doubled cables consist of two cables spooled with two drums on a shared shaft driven by a single motor, as depicted in Fig. 2. Although certain 4-motor planar CDPR geometries can control both translation *and* rotation in the plane, we choose a geometry which largely precludes rotational motion. This choice was made because the benefit of the additional stiffness enabled by the chosen geometry outweighs the lack of CDPR rotation, especially when coupled with the robot arm’s shoulder joint. The CDPR with robot arm is shown in Fig. 2.

2) *Robot Arm*: The robot arm (Fig. 3) is chosen to supplement the planar CDPR with the dexterity to reach around a plant and take photos from a variety of viewpoints. The robot arm was chosen to have 4DoF in a configuration

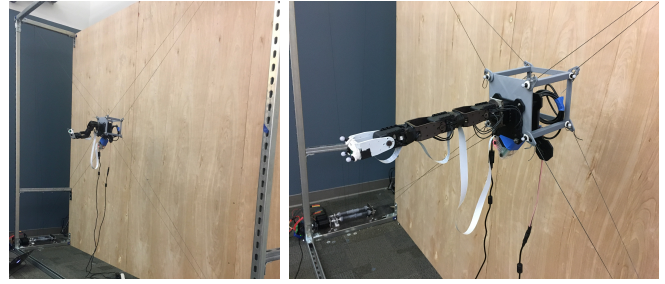


Fig. 2. Left: CDPR consists of 4 pairs of cables controlling a moving platform on which a robot arm is mounted. Right: Each pair of cables is crossed provide additional out-of-plane stability, but both cables in each pair are driven by the same motor.



Fig. 3. 4DoF Robot arm with camera used to take a large number of photos from various angles of a single plant. Left: Arm (without wooden cover or CDPR) taking photos of a plant. Right: Arm inside wooden protective cover.

that, combined with the 2DoF of the CDPR, provides the robot with full SE(3) motion. Although in theory 1DoF is redundant with camera-axis rotation and another 2DoF are unnecessary with a sufficiently wide field-of-view camera, we found in practice that they are helpful when reaching around plants to avoid collisions with neighboring plants.

The robot arm is adapted from a Trossen Robotics PhantomX Pincher Mark II [33], which is a 4DoF robot manipulator using Dynamixel AX-12A servos. The links were extended to have lengths of 0.107 m, 0.194 m, and 0.032 m after the shoulder, elbow, and wrist joints respectively. We also replaced the gripper with a Raspberry Pi Camera Module v2, which uses a IMX219 8MP sensor. The 4 DoF allow rotation in  $\theta$  with the base joint and both translation and rotation in the  $x-r$  plane (see Fig. 4). The completed robot arm is shown in Fig. 3.

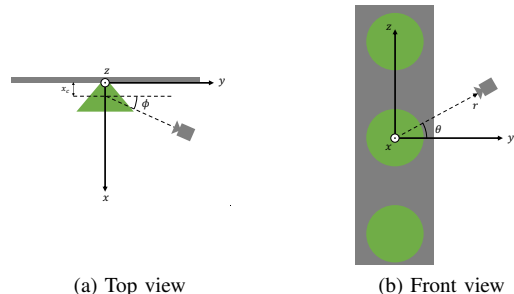


Fig. 4. Coordinate frame of the camera with respect to a lettuce plant.

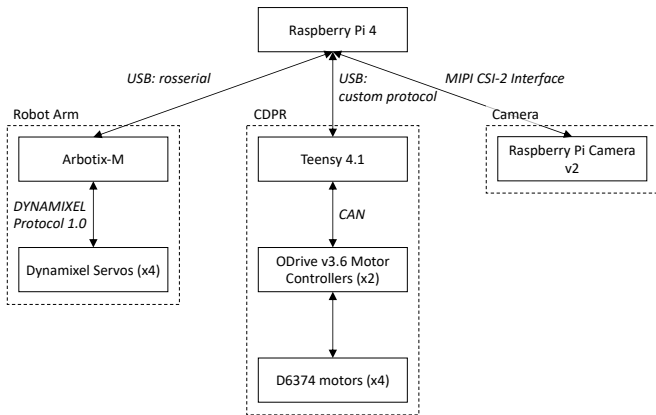


Fig. 5. System communication overview.

### B. Electrical and Communication Design

A Raspberry Pi 4 controls the camera, robot arm, and CDPR using ROS, as overviewed in Fig. 5. The electronics are shown in Fig. 6, excepting the motor controllers and motors which are available in [32, Fig. 6 (right)]. The camera is connected and controlled directly by the Pi. The arm is controlled by an Arbotix-M microcontroller which receives joint angle position commands from the Pi and sends back joint angle position feedback. The CDPR is controlled by a Teensy 4.1 which receives high-level cartesian position commands from the Pi and applies low-level motor torque commands to the motor controllers using the algorithm from [34, Sec. III.A] using factor graphs [35]–[37].

### C. Data Collection Algorithm

The algorithm used for a data collection session consists of using the CDPR to move to a plant, then using the robot arm to take photos from a variety of viewpoints. The positions of the plants on the grow towers are known and pre-programmed for the CDPR to move to, and the set of camera poses is adjusted based on the age of the plant: the view angles remain consistent while the distance to the plant increases with plant age.

## III. EXPERIMENTAL METHODS

We validate our robot design and data collection procedure by collecting data across 2 growth cycles with different nutrient (fertilizer) schedules, analyzing the mass estimates using SfM, and comparing against 3 baseline methods simulating high-throughput or high-accuracy approaches. Although more modern techniques for estimating plant properties using computer vision are available, we use a simple, off-the-shelf SfM approach as it is sufficient to validate the data collection capabilities of our robot.

### A. Lettuce Growing and Dataset Collection Procedure

The data collection procedure was designed to collect photos, masses, and elemental nutrient contents of 72 total plants distributed across various growth stages and 2 different nutrient schedules (“Experiment 1” and “Experiment 2”). We used Bibb Butterhead Lettuce plants because they are

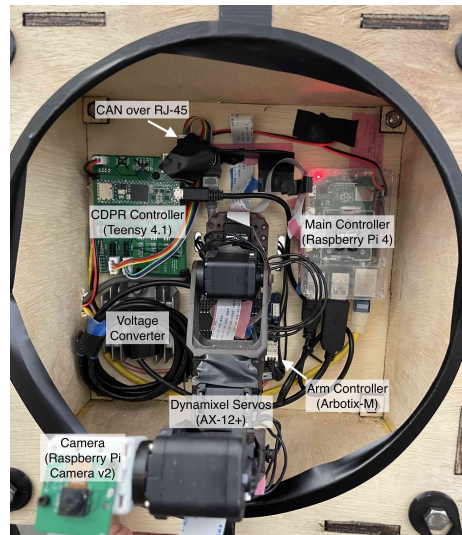


Fig. 6. Electronics mounted on the CDPR moving platform inside a protective wooden cover.

economically well-suited to indoor farming, grow fast, and have qualities that make them challenging for computer vision approaches (e.g. high degree of self-occlusion and self-similar texture).

Due to learned lessons from Experiment 1, the experiments used slightly different procedures, but the growing procedure for each individual plant remained the same:

- 1) To germinate, place seeds in a dampened rockwool substrate, and place the substrate in an incubator next to grow lights for 14 days.
- 2) Transplant the seedlings (with substrate) into the vertical hydroponic grow towers (“Day 0”).
- 3) Take 64 photos (1 top-down, and 21 from each of 3 rings with constant  $\phi$ ) per day.
- 4) Harvest and measure ground truth (GT) data at the scheduled time.

The harvest process consists of cutting the plant at the base, weighing immediately (“Fresh Mass”), dehydrating for 48 hours, weighing again (“Dry Mass”), and performing a nutrient analysis. The fresh mass must be weighed immediately since, once cut, transpiration causes the plant to lose mass so quickly that the reading on a gram scale will observably decrease during the few seconds it is being weighed.

For Experiment 1, the *General Hydroponics Flora Series* fertilizer is used with ratios 3:2:1 of *FloraGro*, *FloraMicro*, and *FloraBloom* totalling 138ml of fertilizer per 100L of water. The pH is monitored and buffered daily, and the hydroponic system is flushed/replenished every 2 weeks. For Experiment 2, the Modified Sonneveld’s solution from [38] is used. The pH is monitored and buffered daily, and the hydroponic system is flushed/replenished 3 times per week.

For Experiment 1, 48 plants were planted in sets of 12 each week and harvested at the same time when the oldest plant reached maturity (28 days after transplant). For Experiment 2, 24 plants were planted at the same time and harvested 3/day from 21 to 28 days after transplant. The

Experiment 2 plants were planted at the same time to reduce variability due to germination conditions. Consequently, they reached maturity at similar times so they were planted at half the density as in Experiment 1 to reduce overcrowding. The harvest schedule was designed to capture the most “interesting” portion of the logistic growth curve.

The specifications for the indoor vertical hydroponic setup will be available in [39], and the vertical grow rig with cable robot is shown in Fig. 1.

### B. Mass Estimation using SfM

We validated the efficacy of our robot and data collection by estimating the masses of the lettuces using SfM. For each plant, we first used COLMAP to generate a dense point cloud of the plant from the photos taken by the robot. Using robot forward kinematics as camera pose priors, we transformed the point cloud into a canonical frame to resolve monocular ambiguity. We then applied a number of programmatic cleaning steps to discard outlier and background points from the point-cloud. Next, we generated a mesh of the plant using Poisson Surface Reconstruction and applied Poisson Disk Sampling to make a more uniform mesh while maintaining a point density sufficient to voxelize without bias on a 3mm grid size. We computed both (1) the surface area of the mesh and (2) a volume estimate by voxelizing the mesh at a 3mm grid size and counting the total occupied voxels. Finally, we applied linear regressions to estimate mass using either surface area or volume.

### C. Baseline Methods

We evaluated our approach against (a) high-throughput methods -satellite, UAV, or conveyor-belt imagery- by using subsets of the total photos collected, and (b) high-accuracy methods by using only the robot arm without the CDPR to replicate similar high-accuracy methods in the literature.

1) *Baseline 1:* Here we used only a single top-down photo of each plant. We applied a 2-layer CNN to segment the plant (foreground) from the background in the undistorted photos and count the number of pixels occupied by the plant. We then used the known pose of the camera relative to the base of the plant, combined with the calibrated camera intrinsics, to approximate the projected area of the plant.

2) *Baseline 2:* Here we simulated over-canopy approaches (such as UAV or conveyor-belt) by using only the photos with camera poses that do not “reach around” the plant. Specifically, we set the minimum threshold for the x-position of the camera to be 17cm (the maximum observed x-dimension of any plant in the datasets) and used only photos from the dataset with camera poses beyond the threshold. We applied the same mass estimation algorithm as in III-B.

3) *Baseline 3:* Here we replicated a low-throughput, high-accuracy method by using only the robot arm without the CDPR, instead manually placing the base of the robot arm in front of each plant. For each plant, we commanded the robot arm to a “top-down photo” pose for reference, fixed the arm’s base in place on a stand (see Fig. 3, left), and took the same set of photos per plant we used for our method. We compared the throughput and camera pose consistency.

TABLE I  
LINEAR REGRESSION RESULTS

Estimation Metric	GT: Fresh Mass		GT: Dry Mass	
	$R^2$	MAE (g)	$R^2$	MAE (g)
Surface Area (ours)	<b>0.845</b>	<b>11.216</b>	<b>0.846</b>	<b>0.586</b>
Volume (ours)	0.833	11.671	0.832	0.617
Baseline 1: Projected Area	0.537	19.976	0.505	1.084
Baseline 2: Surface Area	0.292	26.049	0.285	1.401
Baseline 2: Volume	0.277	26.439	0.269	1.422

## IV. EXPERIMENTAL RESULTS

Using our robot system, we produce a dataset of 54 plants, each containing 64 photos/day, fresh mass, dry mass, and elemental nutrient content. Of the 72 plants initially planned, 18 were discarded due to overcrowding or death. The full datasets are available online [40].

We applied a number of quantitative and qualitative metrics on the datasets to evaluate our design:

- 1) characterize the speed (throughput) of data collection,
- 2) evaluate regressions between GT and estimated metrics (using  $R^2$  and cross-validation MAE),
- 3) compare statistical power using GT vs estimated metrics for hypothesis testing,
- 4) estimate point cloud occlusion using a non-linear regression, and
- 5) visualize point clouds.

### A. Data Collection Throughput

Our robot system was capable of autonomously collecting data at approximately 2640 photos/hour and spanning 56 plants at a density of 350 cm<sup>2</sup>/plant (although the experiments had fewer viable plants than the robot was capable of imaging). Given the inherent scalability of cable robots, increasing the size of the cable robot to reach a greater number of plants should be possible. Additionally, higher quality cameras can dramatically increase the photo capture rate by enabling faster shutter speeds or even continuous robot arm motion (currently, the arm must stop for each photo to eliminate motion blur and rolling shutter effects; motion input-shaping may also improve the capture rate).

Whereas Baseline 3 required a pair of skilled humans to collect data and was only capable of collecting 300-600 photos/hour (64 photos/plant) (depending on the skills of the operators), during experiment 2, our robot system was demonstrated to run autonomously without human supervision across several days. We anticipate that the robot can be run 100% autonomously in future growth cycles.

### B. Mass Estimation Regression Results

Based on  $R^2$  and leave-one-out cross-validation MAE values, our estimated metrics using the full dataset are better correlated to the GT masses than the baseline metrics. Figure 7 shows the linear regressions and  $R^2$  values for a couple representative pairs, and Table I shows the  $R^2$  and MAE values for linear regressions between estimated metrics and ground truth masses.

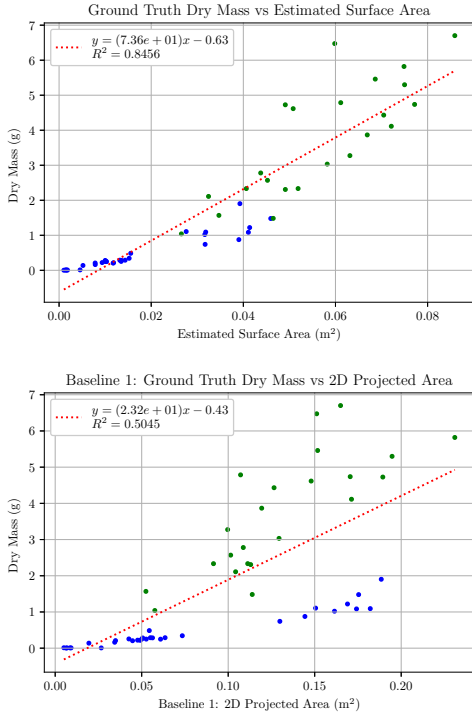


Fig. 7. Sample linear regression results estimating dry mass (GT) using the surface area computed from SfM (top) and the projected area estimate from a single top-down photo of each plant (bottom). Our regression is better than the baseline’s which demonstrates that the our robot design’s ability to capture many viewpoints obtains better performance than a simulated high-throughput approach with limited viewpoints of a plant.

### C. Statistical Power

Ultimately, one primary application of non-destructive phenotyping is to draw scientific conclusions from the computed metrics. To this end, we evaluated our method by running statistical significance tests using (1) ground-truth masses, (2) estimated masses our method in III-B, and (3) estimated masses using the baseline methods from III-C. We used ANOVA [41] to test 2 hypotheses: (a) the age of the plant is correlated with the mass of the plant and (b) nutrient schedule is correlated with the mass of the plant.

To test the hypothesis that age is correlated with mass, we grouped the plants into their harvest days and run one-way ANOVA tests to determine if the different groups have different means of each of the mass metrics. Due to the non-uniform frequencies of ages in the harvested plants, shown in Table II, we performed separate tests for Experiments 1 and 2, and for Experiment 1 we only used the 15 and 21 day-old age groups (note that 2-group ANOVA is equivalent to t-test with independent samples [41]). The results are presented in columns 2 and 3 of Table III and show that both our estimate volume and surface area have similar statistical powers as the GT masses ( $p < 0.005$ ), while the baselines are nearly an order of magnitude less powerful. The exception is Baseline 1 for Experiment 1, which is likely an artifact of the camera position algorithm used for experiment 1 using different camera positions for the different age groups.

TABLE II  
PLANT HARVEST AGE DISTRIBUTION  
# of Samples for Each Age

Age (days)	8	15	21	22	23	24	25	26	27	28
Experiment 1	6	11	11							3
Experiment 2			3	3	2	3	3	3	3	3

TABLE III  
ANOVA STATISTICAL SIGNIFICANCE TESTS

Metric	p-value for Age Discrimination		p-value for Nutrient Schedule Discrimination
	Exp. 1	Exp. 2	
Fresh Mass (GT)	0.00156	0.00037	0.00284
Dry Mass (GT)	0.00137	0.00263	0.00288
Surface Area (ours)	0.00219	0.00352	0.03134
Volume (ours)	0.00204	0.00338	0.03766
Baseline 1: Projected Area	0.00086	0.02661	0.32745
Baseline 2: Surface Area	0.00287	0.31166	0.32066
Baseline 2: Volume	0.00265	0.26535	0.28106

To test the hypothesis that nutrient schedule is correlated with plant mass, we tested whether Experiments 1 and 2 have different mean mass metrics (since Experiments 1 and 2 were executed with different nutrient schedules, as in III-A). To balance the data, only 28 day-old plants were used. The results are presented in column 4 of Table III and show that our method’s metrics are less powerful than the GT masses, but still more powerful than the baseline methods.

### D. Point Cloud Occlusion

We claim that a key advantage of our robot design is that the dexterity of the robot arm allows us to see more of the lettuce plant thereby reducing occlusions (as compared to high-throughput, over-canopy methods). To assess the degree of occlusion, we used a number of assumptions to generate a metric for “occlusion proportion”. Intuitively, we observed that the smallest plants had negligible occlusion so their estimates should be the most accurate. This is reflected by the slightly convex shapes of the data in Figure 7, indicating that extrapolating the lettuce density from small plants would produce under-estimates for the masses of large plants. Assuming that underestimates in the mass are due primarily to occlusion, we approximated the occlusion proportion  $d \approx 1 - \frac{\hat{m}}{m}$  where  $m$  is the true mass,  $\hat{m} := \rho X_{est}$  is a naively estimated mass (because it does not compensate for occlusions), and  $X_{est}$  is the computed metric from the plant photos (surface area, volume, or projected area). We also approximated  $d$  to be proportional to the depth of a cube with the same mass and density as the lettuce:  $d \approx km_{est}^{1/3}$ , where  $k$  (“occlusion coefficient”) indicates a method’s susceptibility to occlusion (smaller  $k$  is better). Finally, we can derive a new equation for the occlusion-compensated mass estimate:

$$\hat{m}_{occ} = \frac{\hat{m}}{1 - d} = \frac{\rho X_{est}}{1 - k(\rho X_{est})^{1/3}}. \quad (1)$$

By running a new regression using this model instead of the linear one in IV-B, we can compare  $k$  for different methods, where small  $k$  indicates less occlusion.

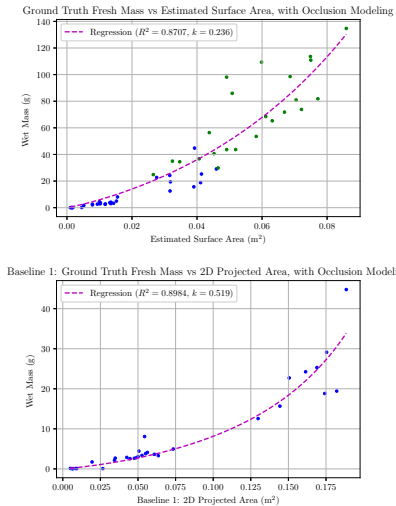


Fig. 8. Representative regression results using the occlusion-compensated model (Eq. (1)) to estimate fresh mass (ground truth) from (top:) the surface area computed from SfM and (bottom:) projected area estimate from a single top-down photo of each plant, suggest that our method is less susceptible to occlusion. Lower  $k$  signifies less occlusion (better).

TABLE IV  
SUSCEPTIBILITY OF DIFFERENT METHODS TO OCCLUSION

Estimation Method	Occlusion coefficient, $k$ ( $\text{g}^{-1}$ ) (lower is better)	
	GT: Fresh Mass	GT: Dry Mass
Surface Area	<b>0.236</b>	<b>0.593</b>
Volume	0.261	0.659
Baseline 1: Projected Area	0.519	0.883
Baseline 2: Surface Area	0.333	0.680
Baseline 2: Volume	0.350	0.743

The occlusion-compensated regression results show that our method has the smallest occlusion coefficient, indicating that the dexterity of the robot arm was helpful in reducing occlusions and capturing a more complete reconstruction of the plant. Representative regressions are shown in Figure 8 and the occlusion coefficients are shown in Table IV.

#### E. Point Cloud Visualizations and Qualitative Descriptions

Both the photos (Fig. 9) and resulting point clouds (Fig. 10) show that our robot is capable of capturing lettuce plant photos from a variety of viewpoints for use in non-destructively estimating plant mass. Fig. 10 identifies gaps in the reconstruction due to occlusion that would result from an over-canopy approach (Baseline 2), thereby supporting the claim that the additional dexterity afforded by our hybrid CDPR with robot arm is helpful in reducing occlusions.

A comparison between photos taken by our robot design and photos taken by Baseline 3 (the arm alone), depicted in Fig. 9, demonstrates the improved consistency with which our robot can position the camera for photos. Camera pose relative to the plant center is highly variable due to human placement error, and the human labor/oversight required is significant (6-9 plants / hour vs no supervision required for the CDPR). 3D reconstructions for Baseline 3 are not shown: aside from the occasional missing points due to improperly framed photos (human error), the reconstructions were similar to those obtained using our robot design.



Fig. 9. Left: Example photos from our plant dataset depict the consistency with which our robot is able to photograph different plants from the same relative camera angle. Right: In contrast, photos taken using only the robot arm (Baseline 3) are inconsistent and labor-intensive.

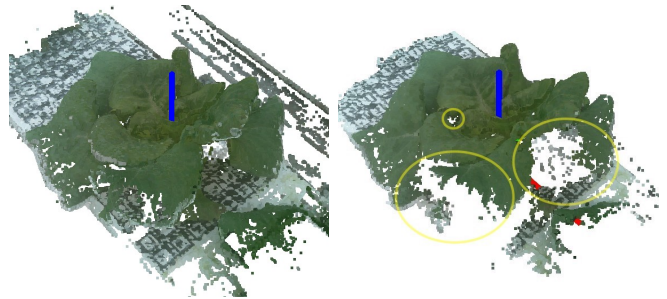


Fig. 10. Dense reconstructions of an example lettuce plant using all the photos (left) vs a subset of photos according to Baseline 2 (right) collected by our robot show that the additional camera viewpoints enabled by our hybrid CDPR+arm design are helpful in reducing occlusions (circled in yellow) and capturing a more complete reconstruction of the plant.

## V. CONCLUSIONS AND FUTURE WORK

In this work, we designed and validated a medium-throughput, high-accuracy plant monitoring robot. We demonstrated its utility by estimating the masses of leafy plants across 2 growth cycles and comparing against 2 baselines simulating higher-throughput systems with fewer available camera viewpoints. We achieved comparable statistical power to ground-truth mass measurements, signifying scientific utility. We also achieved consistently superior performance to the baselines, signifying that, compared to high-throughput approaches to plant monitoring, we achieve higher accuracy in return for our slightly reduced throughput. As compared to prior approaches prioritizing (1) high-throughput but limited viewpoints or (2) low-throughput but many viewpoints, our approach strikes a balance to autonomously collect large numbers of plant photos from a diverse, repeatable set of viewpoints.

Future works include applying visual servo-ing and other real-time computer vision processing techniques to collect better framed photos, analyzing the minimal mechanical design required of the robot arm to achieve high accuracy, and applying more modern techniques for estimating plant properties using computer vision.

## ACKNOWLEDGMENTS

We thank Andrew Sharkey, Thomas Igou, and Teagan Groh for their growing assistance and horticultural expertise.

## REFERENCES

- [1] A. R. Cohen, G. Chen, E. M. Berger, S. Warrier, G. Lan, E. Grubert, F. Dellaert, and Y. Chen, "Dynamically controlled environment agriculture: Integrating machine learning and mechanistic and physiological models for sustainable food cultivation," *ACS ES&T Engineering*, vol. 2, no. 1, pp. 3–19, 2022. [Online]. Available: <https://doi.org/10.1021/acsestengg.1c00269>
- [2] A. Atefi, Y. Ge, S. Pitla, and J. Schnable, "Robotic technologies for high-throughput plant phenotyping: Contemporary reviews and future perspectives." *Front Plant Sci*, vol. 12, p. 611940, 2021.
- [3] M. M. Rahaman, D. Chen, Z. Gillani, C. Klukas, and M. Chen, "Advanced phenotyping and phenotype data analysis for the study of plant growth and development," *Frontiers in plant science*, vol. 6, p. 619, 2015.
- [4] M. Tariq, M. Ahmed, P. Iqbal, Z. Fatima, and S. Ahmad, *Crop Phenotyping*. Singapore: Springer Singapore, 2020, pp. 45–60. [Online]. Available: [https://doi.org/10.1007/978-981-15-4728-7\\_2](https://doi.org/10.1007/978-981-15-4728-7_2)
- [5] L. Li, Q. Zhang, and D. Huang, "A review of imaging techniques for plant phenotyping," *Sensors*, vol. 14, no. 11, pp. 20078–20111, 2014. [Online]. Available: <https://www.mdpi.com/1424-8220/14/11/20078>
- [6] X. Yu, J. Hyypää, M. Karjalainen, K. Nurminen, K. Karila, M. Vastaranta, V. Kankare, H. Kaartinen, M. Holopainen, E. Honkavaara, A. Kukko, A. Jaakkola, X. Liang, Y. Wang, H. Hyypää, and M. Katoh, "Comparison of laser and stereo optical, SAR and InSAR point clouds from air- and space-borne sources in the retrieval of forest inventory attributes," *Remote Sensing*, vol. 7, no. 12, pp. 15933–15954, 2015. [Online]. Available: <https://www.mdpi.com/2072-4292/7/12/15809>
- [7] W.-g. Li, H. Li, and L.-h. Zhao, "Estimating rice yield by HJ-1A satellite images," *Rice Science*, vol. 18, no. 2, pp. 142–147, 2011. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1672630811600206>
- [8] S. Aich, A. Josuttis, I. Ovsyannikov, K. Strueby, I. Ahmed, H. S. Duddu, C. Pozniak, S. Shirliff, and I. Stavness, "DeepWheat: Estimating phenotypic traits from crop images with deep learning," in *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*, 2018, pp. 323–332.
- [9] L. Carlone, J. Dong, S. Fenu, G. Rains, and F. Dellaert, "Towards 4D crop analysis in precision agriculture: Estimating plant height and crown radius over time via expectation-maximization," in *ICRA Workshop on Robotics in Agriculture*, 2015, pp. 1–8.
- [10] Z. Ni, T. F. Burks, and W. S. Lee, "3D reconstruction of plant/tree canopy using monocular and binocular vision," *Journal of Imaging*, vol. 2, no. 4, 2016. [Online]. Available: <https://www.mdpi.com/2313-433X/2/4/28>
- [11] F. Hosoi and K. Omasa, "Voxel-based 3-d modeling of individual trees for estimating leaf area density using high-resolution portable scanning lidar," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 12, pp. 3610–3618, 2006.
- [12] N. Tilly, D. Hoffmeister, Q. Cao, S. Huang, V. Lenz-Wiedemann, Y. Miao, and G. Bareth, "Multitemporal crop surface models: accurate plant height measurement and biomass estimation with terrestrial laser scanning in paddy rice," *Journal of Applied Remote Sensing*, vol. 8, no. 1, p. 083671, 2014. [Online]. Available: <https://doi.org/10.1117/1.JRS.8.083671>
- [13] P. M. Blok, E. J. van Henten, F. K. van Evert, and G. Kootstra, "Image-based size estimation of broccoli heads under varying degrees of occlusion," *Biosystems Engineering*, vol. 208, pp. 213–233, 2021. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1537511021001203>
- [14] E. Marks, F. Magistri, and C. Stachniss, "Precise 3D reconstruction of plants from UAV imagery combining bundle adjustment and template matching," in *2022 International Conference on Robotics and Automation (ICRA)*, 2022, pp. 2259–2265.
- [15] W. Shi, R. van de Zedde, H. Jiang, and G. Kootstra, "Plant-part segmentation using deep learning and multi-view vision," *Biosystems Engineering*, vol. 187, pp. 81–95, 2019. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1537511019308098>
- [16] A. Bannari, D. Morin, F. Bonn, and A. R. Huete, "A review of vegetation indices," *Remote Sensing Reviews*, vol. 13, no. 1–2, pp. 95–120, 1995. [Online]. Available: <https://doi.org/10.1080/02757259509532298>
- [17] D. J. Mulla, "Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps," *Biosystems Engineering*, vol. 114, no. 4, pp. 358–371, 2013, special Issue: Sensing Technologies for Sustainable Agriculture. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1537511012001419>
- [18] J. P. Dandois and E. C. Ellis, "Remote sensing of vegetation structure using computer vision," *Remote Sensing*, vol. 2, no. 4, pp. 1157–1176, 2010. [Online]. Available: <https://www.mdpi.com/2072-4292/2/4/1157>
- [19] J. Geipel, J. Link, and W. Claupein, "Combined spectral and spatial modeling of corn yield based on aerial images and crop surface models acquired with an unmanned aircraft system," *Remote Sensing*, vol. 6, no. 11, pp. 10335–10355, 2014. [Online]. Available: <https://www.mdpi.com/2072-4292/6/11/10335>
- [20] J. Bendig, K. Yu, H. Aasen, A. Bolten, S. Bennertz, J. Broscheit, M. L. Gnyp, and G. Bareth, "Combining UAV-based plant height from crop surface models, visible, and near infrared vegetation indices for biomass monitoring in barley," *International Journal of Applied Earth Observation and Geoinformation*, vol. 39, pp. 79–87, 2015. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0303243415000446>
- [21] D. Li, C. Quan, Z. Song, X. Li, G. Yu, C. Li, and A. Muhammad, "High-throughput plant phenotyping platform (HT3P) as a novel tool for estimating agronomic traits from the lab to the field." *Frontiers in Bioengineering and Biotechnology*, vol. 8, p. 623705, 2020.
- [22] R. Xu and C. Li, "A review of high-throughput field phenotyping systems: Focusing on ground robots," *Plant Phenomics*, vol. 2022, p. 9760269, 2022. [Online]. Available: <https://doi.org/10.34133/2022/9760269>
- [23] J. Lumme, M. Karjalainen, H. Kaartinen, A. Kukko, J. Hyypää, H. Hyypää, A. Jaakkola, and J. Kleemola, "Terrestrial laser scanning of agricultural crops," *The International Achieves of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 47, pp. 563–566, 01 2008.
- [24] G. Van Der Heijden, Y. Song, G. Horgan, G. Polder, A. Dieleman, M. Bink, A. Palloix, F. Van Eeuwijk, and C. Glasbey, "SPICY: towards automated phenotyping of large pepper plants in the greenhouse," *Funct. Plant Biol.*, vol. 39, no. 11, pp. 870–877, 2012.
- [25] G. Polder and J. Hofstee, "Phenotyping large tomato plants in the greenhouse using a 3d light-field camera," *American Society of Agricultural and Biological Engineers Annual International Meeting 2014, ASABE 2014*, vol. 1, pp. 153–159, 01 2014.
- [26] J. Dong, J. G. Burnham, B. Boots, G. Rains, and F. Dellaert, "4D crop monitoring: Spatio-temporal reconstruction for agriculture," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*, 2017, pp. 3878–3885.
- [27] C. Wu, R. Zeng, J. Pan, C. C. L. Wang, and Y.-J. Liu, "Plant phenotyping by deep-learning-based planner for multi-robots," *IEEE Robotics and Automation Letters*, vol. 4, no. 4, pp. 3113–3120, 2019.
- [28] U. Lee, S. Chang, G. A. Putra, H. Kim, and D. H. Kim, "An automated, high-throughput plant phenotyping system using machine learning-based plant segmentation and image analysis," *PLOS ONE*, vol. 13, no. 4, pp. 1–17, 04 2018. [Online]. Available: <https://doi.org/10.1371/journal.pone.0196615>
- [29] J. Demby's, A. Shafiekhani, F. B. Fritschi, and G. N. DeSouza, "Spatio-temporal reconstruction and visualization of plant growth for phenotyping," in *2021 IEEE Symposium Series on Computational Intelligence (SSCI)*, 2021, pp. 1–8.
- [30] N. Kirchgessner, F. Liebisch, K. Yu, J. Pfeifer, M. Friedli, A. Hund, and A. Walter, "The ETH field phenotyping platform FIP: a cable-suspended multi-sensor system," *Functional Plant Biology*, vol. 44, no. 1, pp. 154–168, 2017. [Online]. Available: <https://doi.org/10.1071/FP16165>
- [31] G. Bai, Y. Ge, D. Scoby, B. Leavitt, V. Stoerger, N. Kirchgessner, S. Irmak, G. Graef, J. Schnable, and T. Awada, "NU-Spidercam: A large-scale, cable-driven, integrated sensing and robotic system

- for advanced phenotyping, remote sensing, and agronomic research,” *Computers and Electronics in Agriculture*, vol. 160, pp. 71–81, 2019. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0168169918314170>
- [32] G. Chen, S. Baek, J.-D. Florez, W. Qian, S.-W. Leigh, S. Hutchinson, and F. Dellaert, “GTGraffiti: Spray painting graffiti art from human painting motions with a cable driven parallel robot,” in *2022 International Conference on Robotics and Automation (ICRA)*, 2022, pp. 4065–4072.
- [33] <https://web.archive.org/web/20190610001413/https://www.trossenrobotics.com/p/PhantomX-Pincher-Robot-Arm.aspx>.
- [34] G. Chen, S. Hutchinson, and F. Dellaert, “Locally optimal estimation and control of cable driven parallel robots using time varying linear quadratic gaussian (lqg) control,” [https://gerry-chen.com/publications/Chen22iros\\_cdpr\\_control/Chen22iros\\_cdpr\\_tracking\\_control.pdf](https://gerry-chen.com/publications/Chen22iros_cdpr_control/Chen22iros_cdpr_tracking_control.pdf), 2022.
- [35] F. Dellaert and M. Kaess, “Factor graphs for robot perception,” *Foundations and Trends® in Robotics*, vol. 6, no. 1-2, pp. 1–139, 2017.
- [36] G. Chen and Y. Zhang, “LQR control using factor graphs,” <https://gtsam.org/2019/11/07/lqr-control.html>, Nov 2019.
- [37] S. Yang, G. Chen, Y. Zhang, H. Choset, and F. Dellaert, “Equality constrained linear optimal control with factor graphs,” in *2021 IEEE International Conference on Robotics and Automation (ICRA)*, 2021, pp. 9717–9723.
- [38] N. S. Mattson and C. Peters, “A recipe for hydroponic success,” *Inside Grower*, vol. 2014, no. Jan, pp. 16–19, 2014.
- [39] A. Sharkey and Y. Chen, “To be published: analysis on the effect of potassium starvation on lettuce growth.” This citation will be filled in before camera-ready.
- [40] “Experimental dataset links,” Experiment 1: <https://bit.ly/3RFR32b>, Experiment 2: <https://bit.ly/3xgWGXI>.
- [41] P. Mishra, U. Singh, C. M. Pandey, P. Mishra, and G. Pandey, “Application of Student’s t-test, analysis of variance, and covariance.” *Ann Card Anaesth*, vol. 22, no. 4, pp. 407–411, Oct-Dec 2019.