Non-destructive Plant Analysis with Computer Vision and Robotics

Gerry Chen, Frank Dellaert, Yongsheng Chen

Approvals: __________, __________

I. INTRODUCTION

As environmental conservation is becoming an increasingly important issue, efficient agricultural practices which maximize output while minimizing resource input and environmental impact are imperative to exercise and further develop. For example, according to the EPA, nitrogen fertilizer alone contributes, in the form of N2O, at least 4.1% of total US greenhouse gas emissions [1]. Furthermore, runoff from agricultural practices causes other environmental harm such as algae blooms. Meanwhile, most common current farming practices encourage excess, blanket applications of fertilizer as opposed to the targeted, “smart” applications as part of the growing precision agriculture movement being made possible by recent advances in sensors, data, and AI [2].

Plant growth models are imperative for precision agriculture to predict crop yields under various conditions. In controls terms, the biological plant is the plant of a system, whose dynamics we must model in order to apply optimal control. Therefore, farmers seek larger amounts of real-time, actionable data on their crops while researchers seek larger amounts of higher quality, more comprehensive, more statistically significant data on plant growth to study growth models. Clearly, analysis of plants for their masses, nutrient contents, and other properties is of great interest to both farmers and researchers.

Standard methods for analyzing plant mass and nutrient content are destructive: they require harvesting the plant to weight it, dehydrate it, and send it to a lab for testing. This is not only expensive, but also significantly limits the quantity and statistical significance of data collected. For example, studying plant growth often requires tracking plant metrics as it grows, but destructive analyses make it impossible to measure the same plant multiple times since the first measurement requires killing the plant. Instead, many sets of plants must be grown under identical conditions and periodically harvested for analysis. Robotics and computer vision have been shown to offer significant value by enabling non-destructive methods for analyzing plant properties.

In this work, we propose a novel robotic system for collecting photographic data for non-destructively estimating plant properties and use the robot to collect a dataset of 56 plants. Our robot, depicted in Fig. 1, consists of a robot arm mounted on the moving platform of a cable-driven parallel robot (CDPR) to enable taking photos from a large number of viewpoints across a large number of plants. To evaluate our robot system, we collect a dataset of RGB photos of Buttercrunch Lettuce plants and their corresponding harvest masses (dry and wet). For each plant, we collect roughly 150 photos from various angles. We collect data from 56 plants, harvesting at least 9 plants at a time, 5 times in 1-week intervals to obtain data from different points in the plant growth cycle.

II. PRIOR WORK

A. 3D Reconstruction

The works in this section analyze the 3D structures of plants using 3D reconstructions. It is useful to note that they focus on analyzing leaf area analysis, but show results only on plants with limited leaf overlap (unlike the late-stage lettuce we seek to analyze).

1) RGB: Almost all 3D reconstruction methods (even those discussed later that also use additional sensors) leverage RGB cameras given their ubiquity, low cost, and richness of infor-
from standard 2D RGB cameras. High-throughput is sacrificed for high-resolution from a single viewpoint. Due to their relative ubiquity and ease of acquiring depth information, but both Time of flight (ToF) cameras appear to be popular for their richness of information and relative compactness. They are limited to indoor or greenhouse demonstrations. Also mounts the cameras on a robot arm. Emphasize high-throughput phenotyping and combines RGB and ToF with high-powered flash lighting mounted on a rolling cart. High-throughput is sacrificed for fewer and less controlled plant viewpoints. Finally, while applies leaf segmentation in image space before doing a 3D reconstruction, applies a 3D construction first then applies leaf segmentation. Doesn’t explicitly create a full 3D reconstruction, but segments and reasons about 3D plant properties using image-space depth maps of foreground leaves.

b) ToF cameras: Time of flight (ToF) cameras appear to be popular for their richness of information and relative compactness. All use ToF cameras in combination with one or more RGB cameras to measure leaf size properties. Combining ToF with RGB leverages both the (low resolution) depth information from ToF and the high resolution color information from RGB. They are limited to indoor or greenhouse demonstrations. Also mounts the cameras on a robot arm. Emphasize high-throughput phenotyping and combines RGB and ToF with high-powered flash lighting mounted on a rolling cart. High-throughput is sacrificed for fewer and less controlled plant viewpoints. Finally, while applies leaf segmentation in image space before doing a 3D reconstruction, applies a 3D construction first then applies leaf segmentation. Doesn’t explicitly create a full 3D reconstruction, but segments and reasons about 3D plant properties using image-space depth maps of foreground leaves.

c) Light field (Plenoptic): Although less common, light field cameras have been used as alternatives to ToF or stereo depth cameras for their improved resolution and lack of “unknown distance” pixels due to occlusion (respectively). Light field cameras operate by using special lensing techniques to capture the same image at different depths of field, which can then be analyzed to interpret scene depth information at high resolution from a single viewpoint. Due to their relative rarity, these cameras tend to be expensive despite using parts from standard 2D RGB cameras.

III. APPROACH: ROBOT DESIGN

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 structure from motion 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) Robot Arm: The robot arm is dextrous to allow reaching around a plant to take a large number of photos from a variety of viewpoints. It is adapted from a Trossen Robotics PhantomX Pincher Mark II, which is a 4 degree of freedom (DoF) robot arm with a 31cm reach. The 3D CAD model, from which the kinematic geometry parameters can be ascertained, is available. The robot arm was modified to replace the gripper with a Raspberry Pi Camera Module v2, which uses an 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. 2). The completed robot arm is shown in Fig. 3.

2) CDPR: The CDPR is an 8-cable, 4-motor planar CDPR with a workspace of roughly 2.9m x 2.3m. Details on the CDPR design can be found in [36], 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. 4. The CDPR with robot arm are shown in Fig. 5.

B. Electrical and Communication Design

A Raspberry Pi 4 controls the camera, robot arm, and CDPR using ROS, as overviewed in Fig. 6. The electronics are shown in Fig. 7, excepting the motor controllers and motors which are available in [36, Fig. 6 (right)].

The camera is connected directly to the Pi using a MIPI CSI-2 interface. During visual servo-ing, the camera is used in video mode to process 1080p frames at around 5fps. Upon completion of visual servo-ing, a final high-resolution (2592x1944) photo is taken and saved.

The robot arm is controlled through an Arbotix-M microcontroller which communicates to the Pi using rosserial over USB and to the servos using the DYNAMIXEL Protocol 1.0: a 1-wire, bidirectional, single-master-multiple-slave bus protocol. The robot arm inverse kinematics are implemented analytically and the task-space coordinates are defined in terms of $\theta, \phi, x_c, r$. The Pi issues servo position commands reads back positions to wait until successfully reaching the desired pose before proceeding.

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 Section III.A of [37].
C. Data Collection Algorithm

The algorithm used for a data collection session, which consists of photographing all the plants, is summarized in Algorithm 1. Critical for collecting good quality photos is the visual-servo routine described in Algorithm 2, which is used to center the plants in the photos by panning up/down and zooming out as needed until the entire plant fits is contained in the photo.

IV. Evaluation Procedure

A. Dataset Collection Procedure

The data collection procedure is designed to collect photos, masses, and nutrient contents of approximately 48 total plants distributed across 6 different stages in their growth cycle.

The dataset collection procedure is as follows:

1) Start a new grow cycle for 8 plants each week for 6 weeks.
2) Upon transplanting the youngest set of 8 plants, photograph all the plants according to Algorithm 2, harvest all the plants, and measure the wet mass, dry mass, and nutrient contents (if applicable) of all plants.

Each grow cycle procedure is as follows:

1) Place 12 seeds (Bibb Butterhead Lettuce) in rockwool substrate.
2) Dampen the rockwool substrate with water and place in an incubator next to grow lights for 14hrs/day, as shown in Fig. 8.
3) After 2 weeks, transplant up to 8 successfully germinated seedlings (randomly selected, if applicable) from the incubator to the vertical hydroponic growing towers. The grow tower conditions are under grow lights 14hrs/day; use General Hydroponics Flora Series fertilizer with ratios 3:2:1 of FloraGro, FloraMicro, and FloraBloom totalling 138ml of fertilizer per 100L of water; and are appropriately pH buffered according to the usage directions of General Hydroponics pH Control Kit.

The specifications for the grow lights, densities, and geometries are given in [], and the grow rig with cable robot is shown in Fig. 6.

Algorithm 1 Visual Servo-ing to center the plant in a photo

Require: $\phi, \theta, x_c, r_0$

Ensure: photo $x_c, r \leftarrow x_{c,0}, r_0$

move arm to $(\phi, \theta, x_c, r)$

photo $\leftarrow$ take_photo()

while !is_centered(photo) do

$x, y, \text{size} \leftarrow \text{plant_center_and_size}(\text{photo})$

if $y > \epsilon_y$ or $y < -\epsilon_y$ or $\text{size} > \text{size}_{\text{max}}$ then

zoom out: $r \leftarrow r + \delta_r$

end if

move arm to $(\phi, \theta, x_c, r)$

photo $\leftarrow$ take_photo()

end while

return $\text{photo}, (\phi, \theta, x_c, r)$

Algorithm 2 Photographing of All Plants

for plant index $i = 0$ to $N$ do

move CDPR to plant $i$

for $\phi = \{\phi_0, \phi_1, \ldots, \phi_k\}$ do

for $\theta = \{\theta_0, \theta_1, \ldots, \theta_k\}$ do

photo $\leftarrow$ Algorithm 1 Visual Servo

save [time, photo, $(\phi, \theta, x_c, r)]$

end for

end for

end for

1) Start a new grow cycle for 8 plants each week for 6 weeks.
2) Upon transplanting the youngest set of 8 plants, photograph all the plants according to Algorithm 2, harvest all the plants, and measure the wet mass, dry mass, and nutrient contents (if applicable) of all plants.

Each grow cycle procedure is as follows:

1) Place 12 seeds (Bibb Butterhead Lettuce) in rockwool substrate.
2) Dampen the rockwool substrate with water and place in an incubator next to grow lights for 14hrs/day, as shown in Fig. 8.
3) After 2 weeks, transplant up to 8 successfully germinated seedlings (randomly selected, if applicable) from the incubator to the vertical hydroponic growing towers. The grow tower conditions are under grow lights 14hrs/day; use General Hydroponics Flora Series fertilizer with ratios 3:2:1 of FloraGro, FloraMicro, and FloraBloom totalling 138ml of fertilizer per 100L of water; and are appropriately pH buffered according to the usage directions of General Hydroponics pH Control Kit.

The specifications for the grow lights, densities, and geometries are given in [], and the grow rig with cable robot is shown in Fig. 6.
Fig. 8. 1-week (left) and 0-week (right) old seedlings in the incubator (lid removed) prior to transplantation into the vertical grow towers at 2 weeks.

Fig. 9. The plant grow rig with cable robot, with the oldest plants being 28 days after transplant.

The harvest and measurement procedure for each plant is as follows:

1) Cut the vegetative half of the plant as close to the grow medium/roots as is feasible.
2) Immediately measure the mass of the vegetative half of the plant on a scale to obtain the wet mass of the plant.
3) Dehydrate the plant for 48 hours.
4) Measure the mass of the dehydrated vegetative plant half to obtain the dry mass of the plant.
5) If there is sufficient quantity of sample, send the sample to an external lab for nutrient content testing.

To evaluate the quality of our dataset, we qualitatively compare the quality of the photos to 2 baseline methods: (1) top-down-only photos using an iPhone camera and (2) photos taken using the robot arm only with a human manually moving the arm to each new plant.

V. RESULTS AND DISCUSSION

Our robot system is capable of autonomously collecting data at approximately 2640 photos per hour and spans 56 plants at a density of 54in²/plant. Given the inherent scalability of cable robots, increasing the size of the cable robot to reach a greater number of plants is relatively straightforward while 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).

Using our robot system, we produce a dataset of 56 plants consisting of 150 photos of each plant as well as their ground truth wet masses, dry masses, and elemental nutrient contents. Fig. 10 depicts example photos from the dataset. The full dataset will be available online.

Our dataset evidences the efficacy and utility of our robot data collection platform. The speed and consistency with which photos are taken allows for higher throughput and quality data collection. We can compare the cable robot against 2 other baseline methods tested: (1) single top-down images and (2) an automated robot arm without cable robot. Fig. 11 depicts the results from baseline 1: single top-down images. Although this has a low human labor requirement,

<table>
<thead>
<tr>
<th>Age (days since seed)</th>
<th>Age (days since transplant)</th>
<th># of Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>49</td>
<td>35</td>
<td>9</td>
</tr>
<tr>
<td>42</td>
<td>28</td>
<td>10</td>
</tr>
<tr>
<td>35</td>
<td>21</td>
<td>11</td>
</tr>
<tr>
<td>28</td>
<td>14</td>
<td>12</td>
</tr>
<tr>
<td>21</td>
<td>7</td>
<td>12</td>
</tr>
<tr>
<td>14</td>
<td>0</td>
<td>2</td>
</tr>
</tbody>
</table>
Fig. 11. Photos from baseline 1 of top-down photos from an iPhone camera (prior work by Andrew Sharkey). Only a single photo per plant was taken, limiting the comprehensiveness and predictive power of the data.

![Plant Mass vs Image Foreground Area for 4 Different Nutrient Schedules](image)

**Fig. 12.** Regression result using top-down photos from baseline 1 (prior work by Sushmita Warrier).

the maximum information that can be extracted from this data alone is limited. Prior work analyzing this dataset exhibits poor generalization to different nutrient schedules, with excess of 100% error for certain nutrient schedules, as shown in Fig. 12.

Fig. 13 depicts the results from baseline 2: photos taken by a robot arm but without the cable robot. Because the robot arm alone can only reach a single plant, a human is required to manually place the robot at each plant which not only increases human labor cost, but also introduces additional variability in the data due to imprecise robot placement. Both issues eliminate the possibility of this approach for high-throughput phenotyping in the current laboratory environment.

**VI. CONCLUSIONS AND FUTURE WORKS**

In conclusion, our robotic approach achieves our goals of higher-throughput data collection while also collecting photos from sufficient viewpoints so as to create 3D reconstructions of plants. As compared to baselines of (1) high-throughput but limited viewpoints/images per plant and (2) low-throughput but sufficient viewpoints/images per plant, our approach strikes a balance to autonomously collect large numbers of plant photos from a diverse, repeatable set of viewpoints.

The most immediate future work (within the next month) is to use the data to create plant mass and nutrient content estimates. Additional further works include testing with different nutrient schedules, using additional cameras (including depth, IR, and multi-/hyper-spectral), and increasing the scale of the cable robot to monitor a greater number of plants.

**REFERENCES**


