



Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Original papers

Plant high-throughput phenotyping using photogrammetry and imaging techniques to measure leaf length and rosette area



Nan An^{a,*}, Christine M. Palmer^b, Robert L. Baker^c, R.J. Cody Markelz^b, James Ta^b, Michael F. Covington^b, Julin N. Maloof^b, Stephen M. Welch^a, Cynthia Weinig^c

^a Department of Agronomy, Kansas State University, Manhattan, KS 66506, USA

^b Department of Plant Biology, University of California Davis, Davis, CA 95616, USA

^c Department of Botany, University of Wyoming, Laramie, WY 82071, USA

ARTICLE INFO

Article history:

Received 22 January 2016

Received in revised form 30 March 2016

Accepted 1 April 2016

Available online 5 July 2016

Keywords:

High-throughput

Phenotyping

Photogrammetry

Image processing

Shade avoidance

Arabidopsis thaliana

ABSTRACT

Plant phenotyping is central to understand causal effects of genotypes and environments on trait expression and is a critical factor in expediting plant breeding. Previously, plant phenotypic traits were quantified using invasive, time-consuming, labor-intensive, cost-inefficient, and often destructive manual sampling methods that were also prone to observer error. In recent years, photogrammetry and image processing techniques have been introduced to plant phenotyping, but cost efficiency issues remain when combining these two techniques within large-scale plant phenotyping studies. Using these high-throughput techniques in basic plant biology research and agriculture are still in the developmental stages but show great promise for rapid phenotyping, which will materially aid both science and crop improvement efforts.

In this study, we introduce an automated high-throughput phenotyping pipeline using affordable imaging systems and image processing algorithms to build 2D mosaicked orthophotos. Chamber-based and ground-level field implementations are used to measure phenotypic traits such as leaf length and rosette area in 2D images. Our automated pipeline has cross-platform capabilities and a degree of instrument independence, making it suitable for various situations.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Global crop production and plant biology research are facing a tremendous challenge in that current production rates will be insufficient to meet the demands of the world's population by 2050 (Bongaarts, 2014). Previous studies (Furbank et al., 2009; Reynolds et al., 2009; Tester and Langridge, 2010) showed that traditional breeding programs cannot sufficiently increase annual crop production for the three major cereal crops: rice, maize, and wheat. In the past decade, advances in genetic technology, such as next generation DNA sequencing, have provided new methods to improve plant breeding techniques. With these new techniques, breeders can potentially increase the rate of genetic improvement by molecular breeding (Phillips, 2010).

Many molecular genetic studies have focused on *Arabidopsis thaliana*, an important model system that has been used for identifying plant genes and determining their functions (Arabidopsis Genome Initiative, 2000). These studies have elucidated plant developmental processes and pathways that may generally con-

tribute to yield in diverse crop species. O'Malley and Ecker (2010) reported that homozygous genome-wide knockout lines were available in *A. thaliana*. Weigel and Mott (2009) generated a sequence database of 1001 *A. thaliana* accessions, enabling comparative genomic analyses of yield. Similarly, the genome sequences of many crops, such as rice, maize, wheat, sorghum, and barley, have also been obtained due to the dramatic reduction in sequencing costs in the past few years (Furbank and Tester, 2011). Because of high-throughput genotyping, it is possible to develop large mapping populations and diversity panels for plant breeding (McMullen et al., 2009).

Unfortunately, in contrast to high-throughput genotyping that offers rapid and inexpensive genomic information extraction, conventional plant phenotyping methods are still labor-intensive and cost-inefficient. This greatly limits our ability to quantitatively relate genes to plant growth, environmental adaptation, and yield. Plant phenotyping methods for smaller plants, such as *A. thaliana*, are mainly dependent on intensive manual work for sampling, handling, and measuring plants often invasively, if not fully destructively. Due to this time-consuming process, very few phenotypic measurements can be acquired during the entire growing period (Arvidsson et al., 2011).

* Corresponding author.

E-mail address: an_198317@hotmail.com (N. An).

In the past few years, there has been increased interest in high-throughput phenotyping approaches in controlled indoor environments (Fiorani and Schurr, 2013). These new approaches linking functional genomics, phenomics, and plant breeding are needed to improve both crop production and crop yield stability and also for efficient screening of high-yielding/stress-tolerant varieties (Bolon et al., 2011). Walter et al. (2007) and Jansen et al. (2009) used the GROWSCREEN/FLUORO system to measure chlorophyll and leaf counts. Granier et al. (2006) utilized the PHENOPSIS system to automate the soil water content control for screening soil water deficit responses. Many studies (Furbank and Tester, 2011; Dornbusch et al., 2012; Green et al., 2012; Chen et al., 2014; Dornbusch et al., 2014) have extracted certain phenotypes using LemnaTec Scanalyzer HTS systems (<http://www.lemnatec.com>) that scan plant surfaces with imaging or laser systems to acquire and analyze plant images, or 3D point clouds for extracting certain phenotypic traits. The main advantage of the Scanalyzer HTS is that it is a fully-automated processing pipeline containing image acquisition, storage, management, and processing components, along with some subsequent statistical analyses of the resulting data.

Some larger-scale, fully-automated high-throughput phenotyping facilities have also been deployed in the greenhouses or growth chambers of private sector firms such as Monsanto and Dupont Pioneer and a number of advanced national plant research institutions, such as the Australian Plant Phenomics Facility, the European Plant Phenotyping Network, and USDA. In these installations, robotics, precise environmental control, and remote sensing technologies are used to monitor and assess plant growth and development over time. However, such high-end facilities require budgets far beyond those of most research laboratories and may not be suitable for all situations, such as field environments.

To date, current field phenotyping approaches have mainly focused on automated solutions for data acquisition using platforms that integrate a vehicle, robotics, imaging systems, and sensors. Although this is changing, less work has been directed toward automating data storage, processing, and analysis. Due to these considerations and limitations, high-throughput phenotyping under field conditions has not yet reached its full potential.

Many previous indoor and field studies used imaging systems (cameras or scanners) and invasive sampling methods (excised plant parts) to extract phenotypic traits (Candela et al., 1999; Pérez-Pérez et al., 2002; Cookson et al., 2007; Bylesjö et al.,

2008; Ali et al., 2012; Chitwood et al., 2012). These studies, however, did not take into account the optical distortion generated by imaging system lenses and the perspective distortion created by the angle of view. True distances and areas cannot be determined from a 2D image if either optical distortion or perspective distortion are present, and merely facing the imagers straight down does not fix this problem. In particular, when closely packed, the large number of plants more toward the corners of each frame will be distorted by the perspective viewing angle of the wide-angle lens (Fig. 1). The optical distortion and perspective distortion of the imaging system must therefore be removed before measuring any geometric quantities from a 2D image.

The objectives of this study are to (1) present a low-cost and fully-automated high-throughput imaging-based phenotyping pipeline suitable for both controlled environments and the field, (2) develop novel image processing algorithms for measuring time-series leaf length and 2D rosette area, and (3) model the relationship between rosette area and total leaf expansion.

2. Materials and methods

2.1. Imaging pipeline characteristics and design

The pipeline presented here has three advantages compared to other existing systems: (1) a low-cost imaging system, (2) elements of instrument independency, and (3) cross-platform capability. The first advantage is that off-the-shelf, low-cost digital cameras were used as imaging devices. This technique allows phenotypic traits (e.g., leaf length, rosette area, diurnal plant nastic movements, and plant vegetation conditions) to be extracted and measured directly from images.

The second advantage of this pipeline is a degree of instrument independency. For example, high-level scripts were used to interface with camera-manufacturer-supplied image processing software. Because many camera manufacturers provide similar tools, exchanging cameras becomes mainly a matter of altering those interface scripts. The image analysis algorithms can also be modified based on image features for different image sensors. For example, we successfully integrated a multispectral image sensor in our pipeline on a moving platform with proper modifications for computing vegetation indices.



Fig. 1. Original image with optical distortion and perspective distortion. An original image from indoor environment showing plants before optical distortion and perspective distortion correction. The plants from the corners were seriously distorted and true distances and areas cannot be measured directly from the image.

The third advantage is cross-platform capability. Although the image data acquisition and data transfer methods may vary in different applications, the pipeline has a generic structure so that it can be deployed on different phenotyping platforms in multiple environments with minimal modification. Specifically, the pipeline was deployed on two different imaging platforms: a stationary growth chamber platform and a movable field platform. The novel, generic features enabling the pipeline to operate in these very different environments are outlined next.

As depicted by the flowchart shown in Fig. 2, the pipeline contains five sections. (1) *Image data acquisition* used different platforms in indoor vs. field environments to collect time-series images of plant development. Indoor, stationary imaging systems were designed and mounted on each of six shelves in a growth chamber of University of California, Davis (UCD). At the University of Wyoming (UWY), a mobile imaging system was developed for use in the field. (2) *Image data storage, transmission, and management* involved the use of servers at UCD, UWY, the iPlant Collaborative (<http://www.iplantcollaborative.org/>) and Kansas State University (KSU). For both the chamber and field experiments, images were stored locally and then transmitted via iPlant to KSU. This resulted in three complete backups of all image sets:

one at the origin, one at iPlant, and one at KSU. (3) *Image processing* operations include pre-processing, orthophoto generation, and image segmentation. (4) *Metadata reconciliation* is necessary because metadata generated by different sources (i.e., human-entered data and/or some automated data) may conflict regarding the identity of each image. Reconciliation yields the most accurate pairings of genotype and phenotype data. (5) *Phenotypic trait extraction* includes the machine-vision operations that yield the biological data that comprise the ultimate goal of the system. Python, a high-level scripting language, was used to connect and automate the pipeline sections.

2.2. Imaging platform

2.2.1. Indoor imaging platform

The pipeline development was part of a growth chamber experiment conducted at UC Davis for studying the shade-avoidance response within 8 recombinant inbred line (RIL) populations of *A. thaliana* that have the Col accession as a recurrent parent. A total of 108 Canon Powershot S95 cameras were mounted facing straight down on six shelves—three of which simulated sun and the other three shade (Fig. 3A and B). On each shelf, 18 cameras

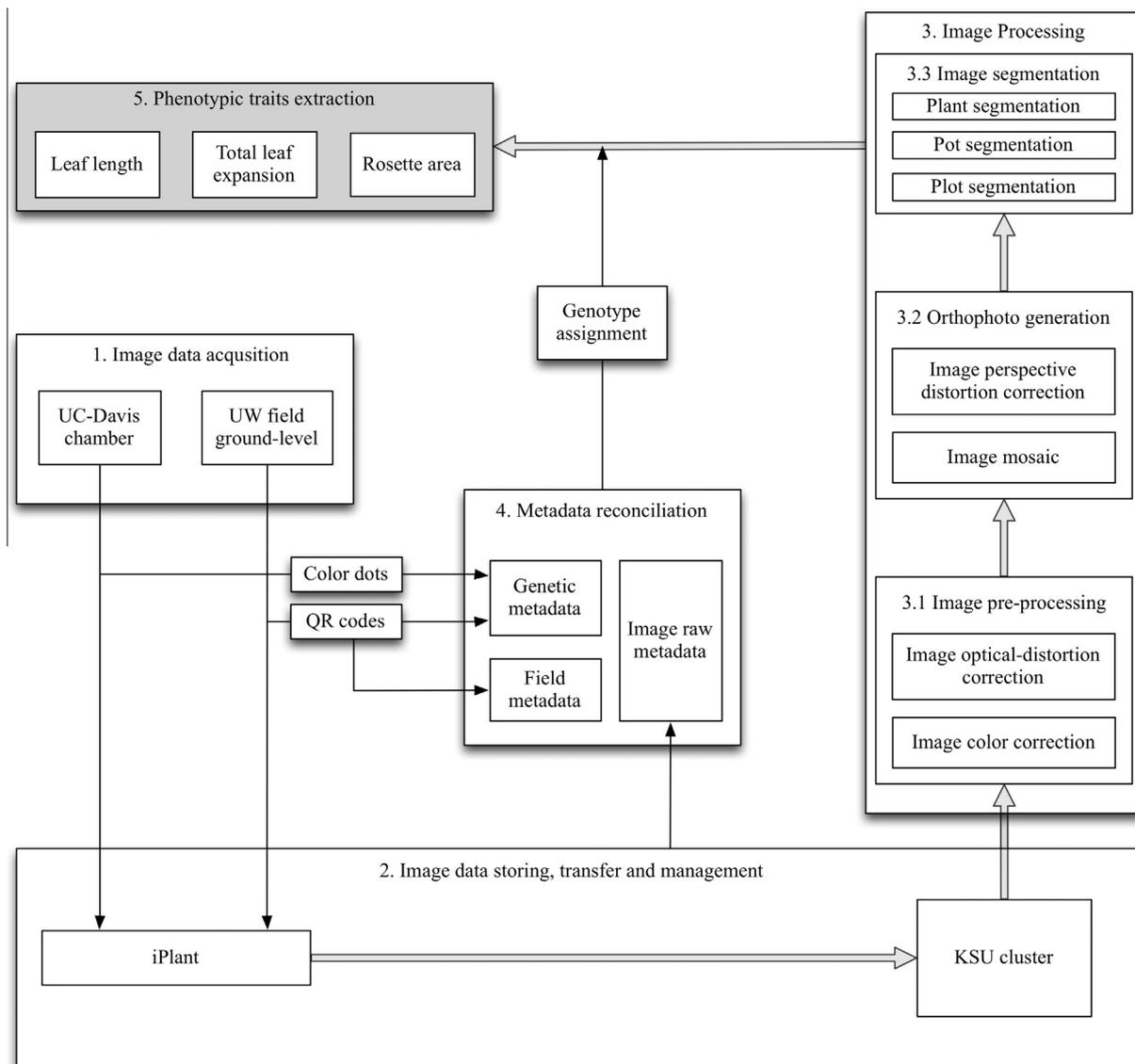


Fig. 2. Pipeline workflow.

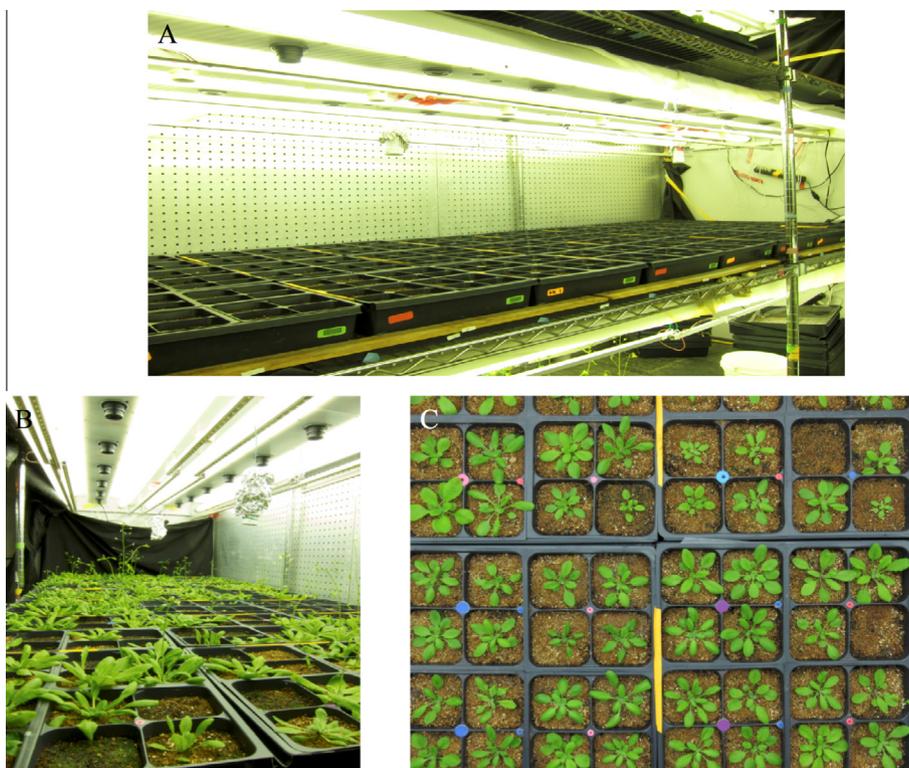


Fig. 3. Indoor imaging platform. (A) and (B) show the chamber imaging system on each shelf. (C) is an individual image from one camera. The color-dot system was used to track 4-by-4 pot flat rotation during the growing period. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

were mounted in a 2-row stationary camera frame 0.4-m above the shelf surface. Each shelf held 24 (three rows of eight) 4-by-4 pot flats within a 0.80-by-2.13-m area. The overall plant density was 225 m^{-1} .

Each camera was assigned a three-digit ID comprised of shelf number (1–6), row number (1–2), and camera position (1–9), in that order. Fig. 3C is one individual image showing the field-of-view (FOV) of each camera and the color dot system for tracking plant rotation. (The color-dot system is described below in the section on indoor genotype assignment.)

All cameras were set on manual focus and exposure modes (F7.1, 1/25 s, auto white balance), and a 28-mm (35-mm equivalent) focal length. The FOV overlap between two adjacent cameras was 50%. A modified intervalometer script and the Canon Hack Development Kit (CHDK) firmware were installed on all cameras to trigger them simultaneously at the start of each hour. In order to prevent camera overheating, the script turned off their LCD screens after each image was taken. All of the images were saved as Canon CR2 RAW format to preserve the maximal amount of geometric, spectral, and camera information.

2.2.2. Field imaging platform

The same genotypes of *A. thaliana* used in the chamber were also planted outdoors at the UWY Plant Science Station in Laramie. Plants were grown for a few days in plastic net pots in the greenhouse and then transplanted to the field in a randomized block design. They were placed in a 10-cm grid with 14 rows and 6 columns. Two Canon EOS REBEL T3i DSLR cameras with Canon EF 20 mm f/2.8 wide-angle lens were mounted on a moveable camera frame at a 95-cm above the ground (Fig. 4). Instead of facing straight down as in the chamber pipeline design, both cameras were mounted angled slightly toward each other to maximize the overlap area of their 80-percent respective FOV.



Fig. 4. Field imaging platform. The ground-level field imaging setup. Two Canon DSLR cameras were mounted with an angle. Two flashes with diffusers were mounted on the side bars for creating a uniform illumination condition.

For stability and repeatability, the camera mount was placed on a metal frame surrounding each plot. The inner frame dimensions were 153.0 cm by 73.6 cm. To image the whole plot, the camera mount was first manually moved to six fixed positions in sequence and pictures were taken. These six pairs did not fully capture all plants, but, because the cameras were on one side of the mount, a seventh position would only have seen the ground outside the plot. Therefore, the camera mount was turned 180° and a final, seventh pair of pictures was taken.

Two Canon flashes with diffusers were attached on opposite sides of the camera mount. These minimized camera mount shadows and the effects of ambient light changes during imaging sessions. Each flash was camera-controlled through an extension cable. Both cameras were set on manual focus and exposure modes (F9, 1/200 s, auto white balance). A customized camera trigger (Industrial Underground Inc., Boulder, CO) was built so both cameras fired simultaneously.

2.3. Image storing, management, and transfer

2.3.1. Indoor image storing, management, and transfer

All cameras were connected to a local UCD data server at via USB cables for transferring images automatically. A Perl script deleted nighttime images (9:00 pm to 4:00 am) and renamed the daytime image files using a combination of the imaging date, time, and camera ID. Although all of the 108 cameras were mounted in landscape orientation, the built-in auto rotation function of the cameras would sometimes rotate images to portrait orientation, creating problems in subsequent steps. To fix this, the Perl-based ExifTool program (<http://www.sno.phy.queensu.ca/~phil/exiftool/>) was integrated into the pipeline to automatically rotate any RAW images found to be in portrait orientation. Each night, the pre-processed RAW images were transferred to iPlant data stores at the University of Texas and, from there, to a server at KSU. Once the images reached the KSU server, they were organized into different subdirectories based on the imaging date, time, and shelf number. This set of transfers resulted in three redundant copies of the images being maintained at UCD, iPlant, and KSU.

2.3.2. Field image storing, management, and transfer

For field image storing and transfer, images were first downloaded from the cameras onto a local computer at UWY and then transferred to the data server at KSU via iPlant. Whereas the indoor system used camera IDs and dates to organize the images, QR codes containing block and plot numbers were employed in the

field (Fig. 5). The QR-code images were first automatically recognized within the stream of images by computing a color histogram and looking for a large number of white pixels. The images containing QR codes were converted to binary using a threshold that removed shadows, then ZBar (<http://zbar.sourceforge.net/>), a free-ware QR coder reader integrated into the pipeline, extracted the block number and plot number. These data were used to group the subsequent images into a directory named by image date, block, and plot information.

2.4. Missing camera detection mechanism of indoor imaging pipeline

During the imaging period, occasionally some cameras would accidentally turn off, possibly due to unstable CHDK firmware. If not immediately detected and corrected, gaps in phenotypic data would result. Therefore, we included in the pipeline a mechanism for detecting missing cameras based on tracking the cameras' IDs in the image names. When missing camera IDs were detected, the pipeline automatically sent an email reporting the problem so it could be manually fixed.

2.5. Pipeline control

Agisoft Photoscan Pro (Photoscan), further described below, includes a Python scripting application program interface (API) whose intent is to allow users to automate its capabilities. This was exploited to control all pipeline functions, including, in some cases, programs completely external to Photoscan. The sequence of pipeline operations is described in the following sections.

2.6. Image pre-processing

There were two corrections performed during the image pre-processing section: color correction and image optical distortion correction. Both corrections as well as conversion to the TIFF file format were performed by the Canon Digital Photo Professional (DPP) program.

2.6.1. Image color correction

Due to illumination variation across the shelf, the camera color responses differed slightly. For better plant segmentation processing in the following step, the DPP white balance operation was used to correct RAW image colors. Enabling this was the reason that the spectrally-lossless RAW file format despite its large memory requirement. Wide-angle lenses are also susceptible to



Fig. 5. Field image data management. Three successive images from one camera. Left to right they are: the last image from the preceding plot, the first image from the next plot with a QR code held in the camera view, and the same plot view with the QR code removed. The image containing the QR code is recognized by its large number of white pixels. The QR code is then read to identify the new plot and the block that contains it.

vignetting effects that reduce image brightness at the periphery. This can complicate color segmentation but was also corrected during this process. Color correction effectiveness was verified by photographing a customized color-grid poster.

2.6.2. Image optical distortion correction

Few previous image-based phenotyping studies have considered lens distortion when extracting leaf shape parameters (length, width, and area). This omission reduces measurement accuracy for plants not at the image centers. The RAW file format also enabled linkage to manufacturer-provided lens data for correcting plant geometric distortion—another function built into DPP. After image color correction and optical distortion correction the memory-intensive RAW format was no longer needed so TIFF image files were exported. The color-grid poster was also used for verifying the image optical distortion correction.

2.6.3. DPP automation

A design drawback of DPP is that it assumes a human will be using it to correct a small number of images. Thus, it lacks any automation capabilities. Therefore, AutoIt (<https://www.autoitscript.com/site/autoit/>), a BASIC-like scripting language, was used to automate the DPP graphical user interface (GUI). This language simulates user mouse clicks and text entries. While this may seem cumbersome, it is a major advantage of the pipeline. Aside from the Perl script described, the AutoIt script is the only element of the pipeline that

would have to be altered if a different brand of camera and manufacturer-provided image correction software were adopted.

2.7. Orthophoto generation

The final type of correction removed image perspective distortion. This was done by generating orthophotos, which are synthetic images produced as if each pixel is being viewed straight down. Thus, orthophotos permit geometric quantities such as 2D distances and areas to be measured with perspective effects removed. The Agisoft Photoscan Pro (Photoscan) program (<http://www.agisoft.com>) performed this step using the TIFF images output from DPP. This was done in nine-image subsets, each of which covered one half-shelf. (The original intent was to do full shelves but it was discovered after plants were added to the chamber that the vertical camera spacing did not permit this—a design flaw to be avoided in the future.) The program converted each set to an orthophoto. However, in the process of implementing this step, a subtle difference between the chamber and the field was uncovered that affected the best way to do this. This is described in the following two subsections.

2.7.1. Indoor environment image rendering method

Photoscan has four alternative rendering options for producing orthophotos: Mosaic, Average, Max Intensity, and Min Intensity. These govern the coloring method used to merge corresponding

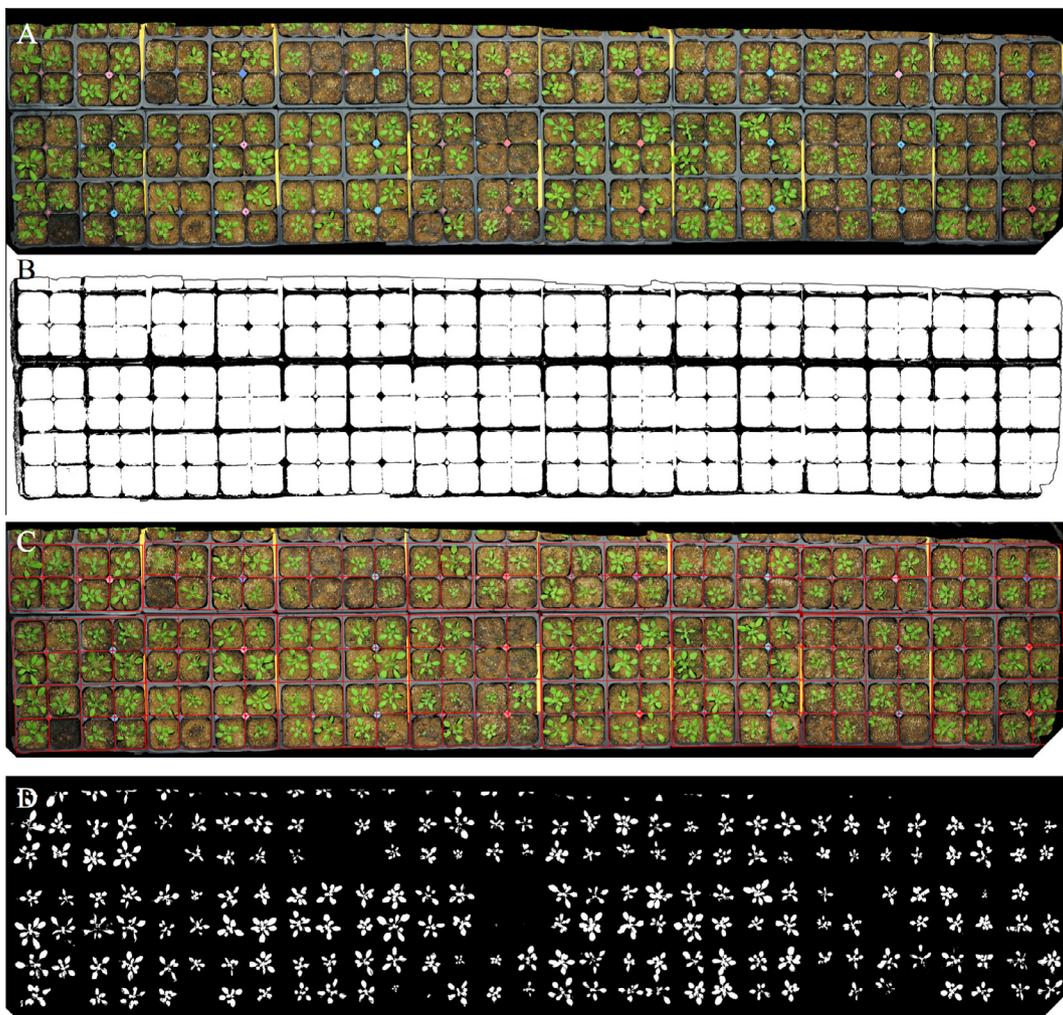


Fig. 6. Orthophoto processing. (A) The mosaicked orthophoto for half-shelf; (B) detected pot binary image; (C) generated 4-by-4 grid overlaid on the orthophoto; (D) detected plant binary image.

pixels from different images into the orthophoto. It was discovered that a wrong choice could have side effects for the small fraction of leaves having greatly different orientations as seen by separate cameras. Specifically, the leaves would appear to be ghost-like double exposures. This was corrected by choosing the Mosaic method that favored the camera whose view of ghost leaves was most vertical. An example orthophoto is shown as Fig. 6A.

2.7.2. Field image rendering method

When used in the field, the Mosaic rendering method left pronounced shadows in the orthophoto that complicated subsequent processing steps. This was resolved by using the Average rendering method. This produces a more uniform orthophoto because areas that were shadowed by the camera mount in one image would often be sunlit in others. Because it blends pixels, the Averaging method reduces shadow contrast.

To summarize the pipeline control description, Python scripts written and executed within Photoscan first invoke AutoIt to run a script in that language simulating user keystrokes and mouse clicks instructing DPP to remove lens and color distortion and produce TIFF formatted images. Once AutoIt processing is finished, the Python script then initiates Photoscan operations that produce the orthophoto. The same script then continues, executing the operations described in the following sections.

2.8. Image segmentation

2.8.1. Pot segmentation

In order to extract individual plants from each chamber orthophoto, the first step is to apply image segmentation to identify the pots. This was done using the following equation:

$$Pot = bw(Blue - Red),$$

where *Blue* and *Red* are the respective blue and red channel pixel brightness values, and *bw()* is the Otsu threshold method (Otsu, 1975) for binary image transformation. Due to the slight illumination variation across each shelf, some of the pot edges could not be detected. The probabilistic Hough Transformation (Duda and Hart, 1972) was implemented in Python to identify the line segments and fill in the missing pot edges (Fig. 6B). A 4-by-4 grid was generated for the pots in each flat based on their known dimensions (Fig. 6C).

2.8.2. Plant segmentation

Plant segmentation, the process of removing unwanted image features like soil, pots, or other items, is the next process applied to each orthophoto. The well-controlled illumination sources and image color corrections in the previous step enabled a simple vegetation index to be used for quick plant segmentation. The Normalized Green–Red Difference Index (NGRDI; Hunt et al., 2005) is similar to the well-known Normalized Difference Vegetation Index (NDVI). However, NDRDI is more useful to distinguish healthy vegetation from background in cameras that have not been modified to be infrared-sensitive. The Otsu threshold method then transformed grayscale NGRDI images to a binary form in which the plant pixels were white and all non-plant pixels were black. The NDRDI equation in this study is as follows:

$$Plant = bw\left(\frac{Green - Red}{Green + Red}\right),$$

where *Green* is the green channel pixel brightness value and *Red* and *bw()* are defined above. Fig. 6D shows processed binary orthophoto. The NGRDI equation was implemented in Python and the Otsu threshold was from the Open Source Computer Vision Library (OpenCV; <http://opencv.org/>).

2.9. Plant genotype assignment

2.9.1. Indoor pipeline genotype assignment

Although growth chambers are well controlled, there are still temperature and lighting gradients that can affect plant growth and development. It is therefore common practice in *Arabidopsis* experiments to randomly reshuffle flats of pots every two to three days. Flat movements are recorded in spreadsheet form, but, to provide redundancy within the image data, each flat was assigned a unique three-color dot combination. A color-dot detection and decoding routine was integrated into the processing pipeline for automatically tracking flats so that the proper genotypes of each plant could be paired with the ultimate measured phenotypes.

2.9.2. Field pipeline genotype assignment

Because plants were arranged in a 14-by-6 grid within each field study plot, a Python routine in the pipeline generates a like-sized grid on each orthophoto for plant extraction. However, it was discovered that plot orthophotos needed to be cropped first so that the extraction grids were properly placed despite some

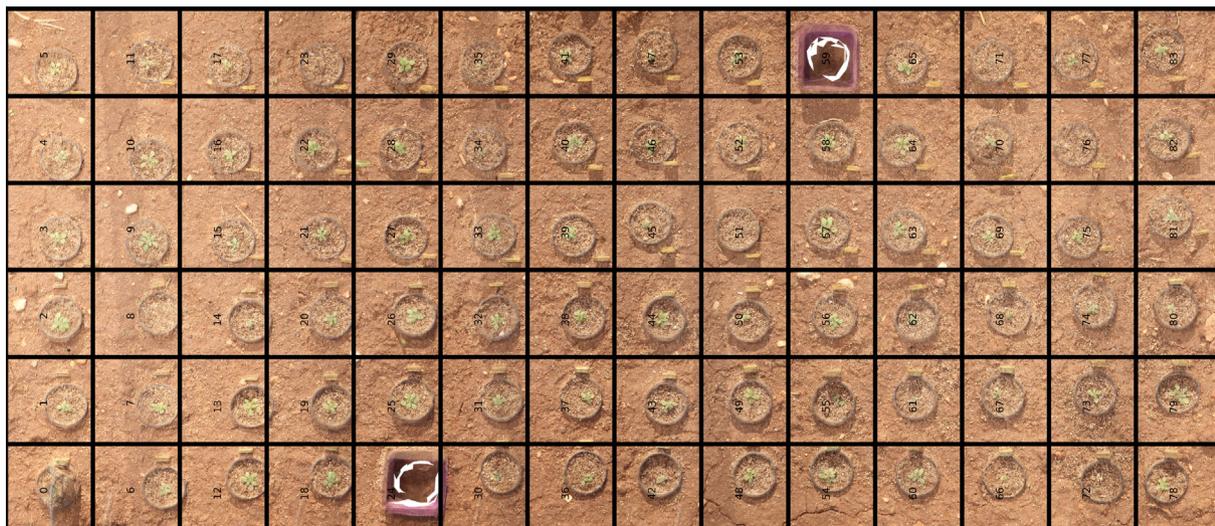


Fig. 7. Field single plant extraction and genotype assignment. The orthophoto of a plot with an automatically-generated black grid for single-plant extraction and genotype assignment.



Fig. 8. Image optical distortion correction. (A) The image before the image optical distortion correction. The red reference line on the top shows the curvature of the edge of the color grid. (B) The image after the image optical distortion correction. The edge of the color grid shows straight comparing the red reference line. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

irregularities in the planting grid placement within each plot (Fig. 7). Each plant was then extracted into a single image named according to plant position and QR-coded block and plot numbers. This information was paired with the genotype metadata collected when plant locations were assigned.

2.10. Phenotypic traits extraction

2.10.1. Leaf length and total leaf expansion calculation

The key step in measuring leaf length is the detection of the rosette center and leaf tips on each single-rosette binary image. The contours of each binary image were analyzed first and the image moments were then calculated (https://en.wikipedia.org/wiki/Image_moment). The rosette center was estimated using the binary image centroid of all white (i.e., plant) pixels. That is:

$$(\bar{x}, \bar{y}) = \left(\frac{M_{10}}{M_{00}}, \frac{M_{01}}{M_{00}} \right),$$

where \bar{x} and \bar{y} are the coordinates of the binary image centroid and M_{ij} are image moments.

Using the calculated rosette center as the origin, a radial scan was executed on the binary image to yield a curve representing the traced rosette outline in a 2D plot. The leaf tips, being the points most distant from the plant center, should be the peaks of the curve just described. However, at first it was challenging to find accurate peak locations due to leaf edge roughness. This was even more complicated when parts of the leaves appeared to be missing due to damage or segmentation faults. Therefore, the rosette-outline curve was first smoothed with a Savitzky–Golay filter (Savitzky and Golay, 1964) to remove small, erroneous maxima.



Fig. 9. Image color correction. (A) The image before the image color correction. The green tone of the image was due to shelf illumination. (B) The image after the color correction shows that the green tone was removed. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

The next step was to fit a Chebyshev polynomial (Tchebychev, 1854) to the smoothed rosette-outline curve. Putative peaks were then located by calculating the roots (i.e., zeros) of the first derivative of the Chebyshev polynomial curve. This procedure still yielded false leaf tip positions occasionally.

Therefore, as a third step, the peak widths were analyzed using the x coordinates of the rosette-outline curve to find the minimum peak width, which was used as the length of a moving window centered at each detected curve peak. Within this moving window, the maximum of the fitted rosette-outline curve and the maximum of the original rosette-outline curve were compared. The true peak locations were recovered if the maximum of the original rosette-outline curve was higher. The pixel coordinates of the leaf tips were calculated based on the curve peak locations. The length of each leaf was measured as the distance from the rosette center to the leaf tip. The “total leaf expansion” of each plant was defined as the sum of all of the leaf lengths.

2.10.2. Rosette area calculation

The rosette area in each single-rosette binary image can be simply calculated by computing the total number of white pixels in each single-rosette binary image.

2.11. Statistical modeling

Chitwood et al. (2012) manipulated far-red light to induce changes in leaf length as an index of the shade-avoidance response. This study demonstrated a linear relationship between total leaf length and square root of the total leaf area. To test the relationship between total leaf expansion and rosette area using our workflow, we fit the phenotype data with the following equation as part of the pipeline:

$$\text{Total Leaf Expansion} = a * \text{Rosette Area}^b,$$

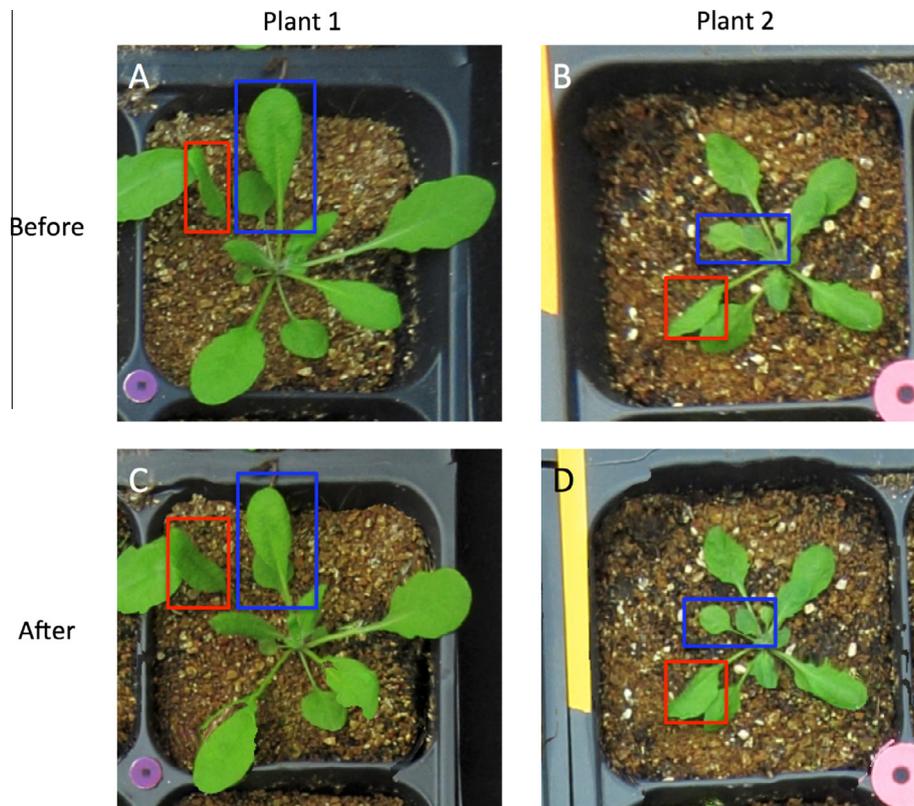


Fig. 10. Image perspective correction. (A) and (B) were plant 1 and plant 2, respectively, before the perspective distortion correction, and (C) and (D) were the same two plants after the correction. The perspective distortion correction can: (1) show more leaf area when that leaf was not flat (as red-box-highlighted), and (2) show the corrected leaf positions (as blue-box-highlighted). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

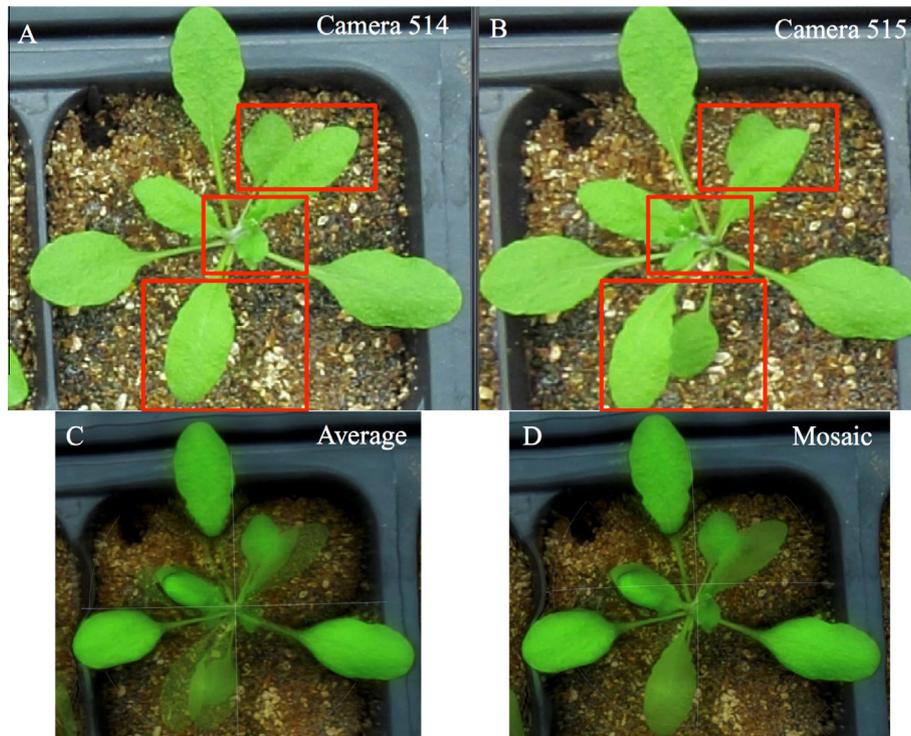


Fig. 11. Indoor orthophoto-rendering methods comparison. (A) and (B) show the same plant photographed by two adjacent cameras. The leaf positions were not consistent due to different viewing angle. When the Average rendering method was used, the “ghost-leaves” issue confused segmentation algorithm, as in (C). The Mosaic blending method could provide better image for segmentation, as in (D).

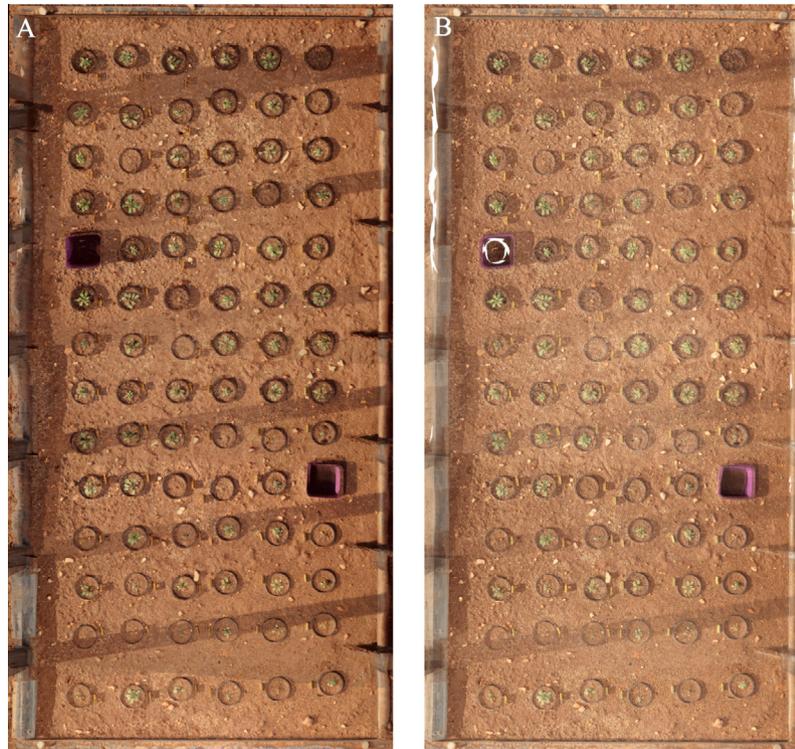


Fig. 12. Field orthophoto-rendering methods comparison. (A) is the ground-level plot orthophoto using Mosaic blending method; (B) is the orthophoto for the same plot using Average blending method. Notice that the camera mount shadows are minimized in B.

where a and b determine the trajectory and shape of the power law function, respectively.

The a and b parameters were estimated using the Levenberg–Marquardt non-linear least square method. Bootstrap resampling was used to calculate their 95% confidence intervals (CI) based on 10,000 simulations. The power law function and the least square estimation were implemented using Python Scipy package.

2.11.1. Indoor pipeline data analysis

Data for total leaf expansion and rosette area collected at 8, 10, 12, and 14 days after plant emergence were analyzed. These time points were selected so that: (1) the plants were big enough to distinguish individual leaves from the start, (2) the final image had a large rosette, but (3) final leaf overlaps between adjacent plants were minimal. All of the images from the indoor pipeline were taken at 6:00 am to minimize any possible ambient influences.

2.11.2. Field pipeline data analysis

The genotypes examined in the indoor environment were also tested in the field. The same power law was fit to data collected on six different days. They were selected from an 18-day period in which Day 1 was the third day after transplantation from the greenhouse. This was followed by five subsequent dates, each three to four days apart as determined by weather conditions.

3. Results and discussion

3.1. Indoor imaging pipeline throughput capability

Because there was a certain amount of unavoidable manual work during image acquisition in the field, overall throughput capability was evaluated using the completely-automated indoor pipeline.

Orthophoto generation, which initially required 38 CPU-minutes per half-shelf, was the most time-consuming process in

the entire pipeline. By implementing High Performance Computing (HPC) routines using OpenCL and AMD GPUs in the pipeline, this processing time was reduced to 25 min per half-shelf orthophoto. These HPC routines were also utilized in other steps. It took approximately six to eight minutes to detect pots and decode the color dots for genotype assignment, then two minutes for extracting phenotypic trait data (e.g., total leaf expansion, rosette area). Therefore, the total runtime of the entire processing pipeline was 35 min maximum for each half-shelf.

The second method for improving pipeline throughput was to use distributed parallel computing to spread the processing tasks across a small computing cluster. Two split half-shelf orthophotos from each shelf were processed by two nodes simultaneously. For the chamber application, there were 24 4-by-4 pot flats on each of the six shelves, which made 2304 plants photographed hourly. Because imaging was conducted for 16 h per day, a total of 36,864 single-plant pictures were obtained each day. Each cluster node processed 192 plant images in 35 min, equating to 10.94 s per plant on average. Therefore, the hourly throughput is 329 plants per hour per cluster node and the runtime for processing 16 h of single-rosette images on the six-node cluster is 11.2 h.

This pipeline was fast because all 108 cameras took images for six shelves simultaneously. The times for image storing, management, and transfer have not been included in this analysis because of variation in local network and internet speeds. In the future, when the processing pipeline is executed on a local computer cluster at UCD for minimizing data transfer time, all 2304 plants can be screened within an hour.

3.2. Image analysis

3.2.1. Image optical distortion and color correction

The 28-mm focal length caused the optical distortion to be quite visible at the image periphery. Fig. 8A shows the color-grid image before optical distortion removal. The straight red reference line

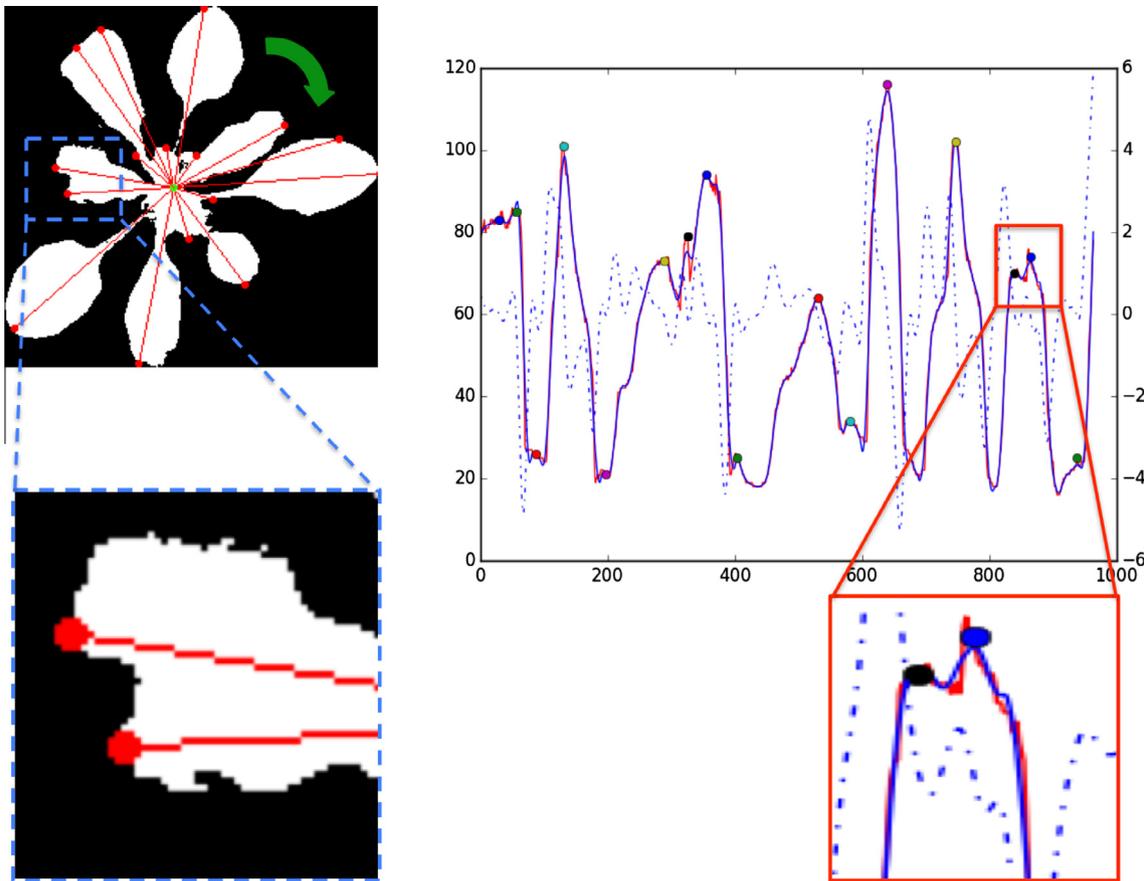


Fig. 13. First iteration of leaf tip detection. The first iteration of the Chebyshev Polynomial fitting for detecting the peaks of the rosette-outline curve, then the corresponding leaf tips on the single-rosette binary image can be found. In the plot, the red curve is the original outline curve, the blue curve is the Chebyshev Polynomial fitting curve, and the blue-dashed curve is the first derivative of the fitted curve. A damaged leaf (as pictured zoomed-in) can show false tips due to the roughness or the damage of leaf edges. The corresponding plant outline curve and peaks are also highlighted. The green-curved arrow shows the direction of the radial scan. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

drawn on the image reveals the color grid curvature caused by the optical distortion. The restored alignment of the red reference in Fig. 8B shows that using Canon DPP and the lens profile database was accurate and efficient at removing such distortion.

Due to variation in shelf illumination and camera firmware differences, the original images were greener than the original RGB values of color grids. After comparing different image color correction packages, the Canon DPP software provided reasonably accurate color correction via its white balance function. After color correction, the green tone was removed and the photographed RGB values were very close to the original (Fig. 9).

3.2.2. Image perspective distortion correction

Fig. 10A and B shows two plants at the corner of an individual frame taken from extreme corners of the shelf where both factors combine to generate perspective distortion. The effects of perspective are clear because a great deal of pot wall can be seen compared to plants at the image centers, where less wall is visible. Due to this issue, direct measurements from the image centers and corners are not comparable; Fig. 10C and D shows that after re-projecting each pixel vertically. The sidewalls of the pots were largely corrected compared to the uncorrected images. Moreover, the red-box-highlighted leaves showed more leaf area in the corrected image for both of the plants when the leaves were not entirely flat.

Moreover, the blue-box-highlighted leaves show that leaf positions were also corrected. For plant 1 (Fig. 10A), the two leaves in the blue box showed a side-by-side position, but in reality, the

smaller leaf was under the big leaf, as the corrected image shows (Fig. 10C). For plant 2, the two leaves highlighted by blue boxes overlapped in the uncorrected image but were distinguished clearly after correction (Fig. 10B and D). Although smaller leaves could not be counted when covered by bigger leaves, the perspective-distortion-corrected images will provide more accurate leaf length and rosette area measurements, which are more critical when studying leaf shade-avoidance responses.

3.2.3. Rendering methods for indoor and field orthophoto

Most previous studies did not use orthophotos for quantifying phenotypic traits from 2D images and therefore did not have occasion to compare rendering methods and the defects they can occasionally introduce. Fig. 11 illustrates the double image “ghost leaves” problem described earlier and how it was resolved using the Mosaic method, which colors¹ the orthophoto using the image pixels that resolve as being closest in 3D space.

The red-box-highlighted leaves in Fig. 11A and B were photographed with different viewing angles by adjacent cameras 514 and 515. In this instance, a small leaf was covered by a bigger leaf. Camera 515 photographed them at an oblique angle so both leaves were visible to this camera but not to Camera 514. Because of this, under the Average rendering method, which blends corresponding pixel colors from both images, the highlighted leaf appears twice in

¹ For interpretation of color in Fig. 11, the reader is referred to the web version of this article.

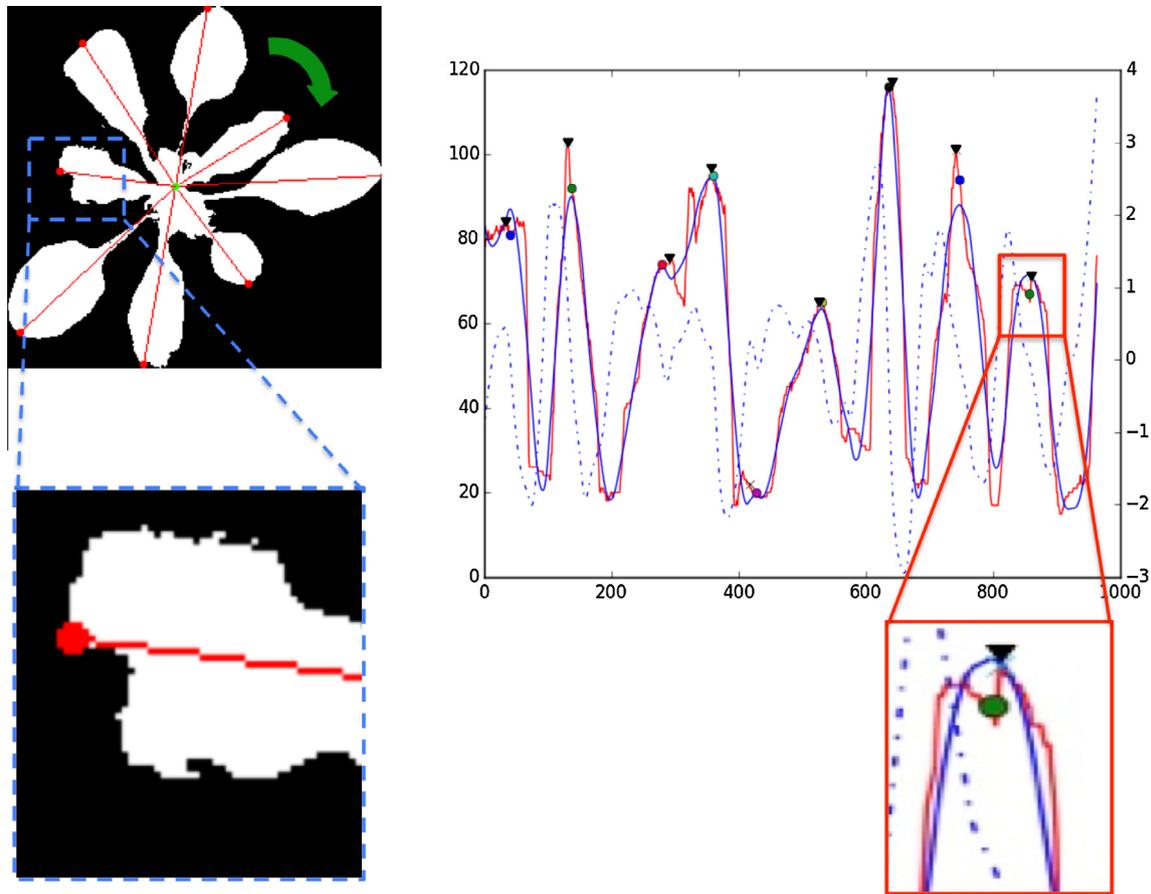


Fig. 14. Second iteration of leaf tip detection. The second iteration for running the optimization algorithms. A smoother curve was fitted to the original outline curve and then a moving window was centered at each peak to locate the highest peak (black triangle markers on the curve) in the window. The zoomed-in view of the curve shows that a better leaf tip is detected. The zoomed-in view of the curve shows that the corresponding peak of the original outline curve (red curve) is detected at the black triangle marker. The true leaf tips could be relocated on the binary image. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

the orthophoto (Fig. 11C). This defect would subsequently confuse the leaf tip detection algorithm. However, in Fig. 11D, when the Mosaic rendering method was used, the plant outline was not confusing and the “double-leaf” issue disappeared.

There are actually three defects in Fig. 11 which need to be discussed: one is partial overlap (top red box), one is different viewing angle (middle red box), and the last is complete overlap (bottom red box). With the Average method, all three create spurious leaf tips due to differences in the “opacity” of the double images. However, the Mosaic method fixes this issue at the cost of occasionally losing an entire leaf (bottom red box).

On the other hand, in the field environment, the two cameras were relatively far away from the plants, so the viewing-angle issue was not as pronounced as in the chamber. Instead, the major problem affecting rendering was strong ambient illumination by sunlight, which cast constantly-changing shadows of the moving camera mounts into the plots. Originally, it was thought that the flashes could be used to completely eliminate shadows. Unfortunately, this required very high intensity, which saturated image brightness. Therefore, a lower setting was used in combination with the Average rendering method to improve the consistency of orthophoto brightness. Fig. 12A shows pronounced camera mount shadows when the Mosaic rendering method was used. Subsequent image segmenting by color analysis was not able to achieve equal results at identifying plants under shadowed and non-shadowed conditions. However, in Fig. 12B, the orthophoto produced by the Average rendering method shows more uniform brightness throughout the entire plot. These images could be suc-

cessfully segmented even though the shadows were not completely eliminated.

3.2.4. Field QR codes imaging and processing

QR codes provided a very efficient way to store plot metadata and organize plot images. Initially, the ground-level field pipeline could successfully recognize all of the QR-code images among other field images by analyzing the histograms of the images. However, despite easy recognition, the Zbar reader sometimes failed to extract the data when camera mount shadows fell on the QR codes. To solve this issue, the brightness and contrast of all QR-code images were first increased to minimize the shadow and then thresholded to produce a shadow-free binary image. With this improvement, all QR codes were successfully decoded and the plot metadata accurately extracted.

3.2.5. Leaf tip detection refinement

Fig. 13 illustrates the irregularities leading to the initial detection of false tips (described above) that compromise accurate length and total leaf expansion measurements. However, the addition of tip windowing to Savitzky–Golay filtering and Chebyshev Polynomial fitting (Fig. 14) resulted in the shifting of peak locations back to the original radial-scan-leaf curve peak locations, the recovery of true peak locations, and the elimination of many false peaks. These optimization operations provide much more accurate rosette-outline peaks for positioning the true leaf tips on the single-rosette binary image. The method was especially accurate

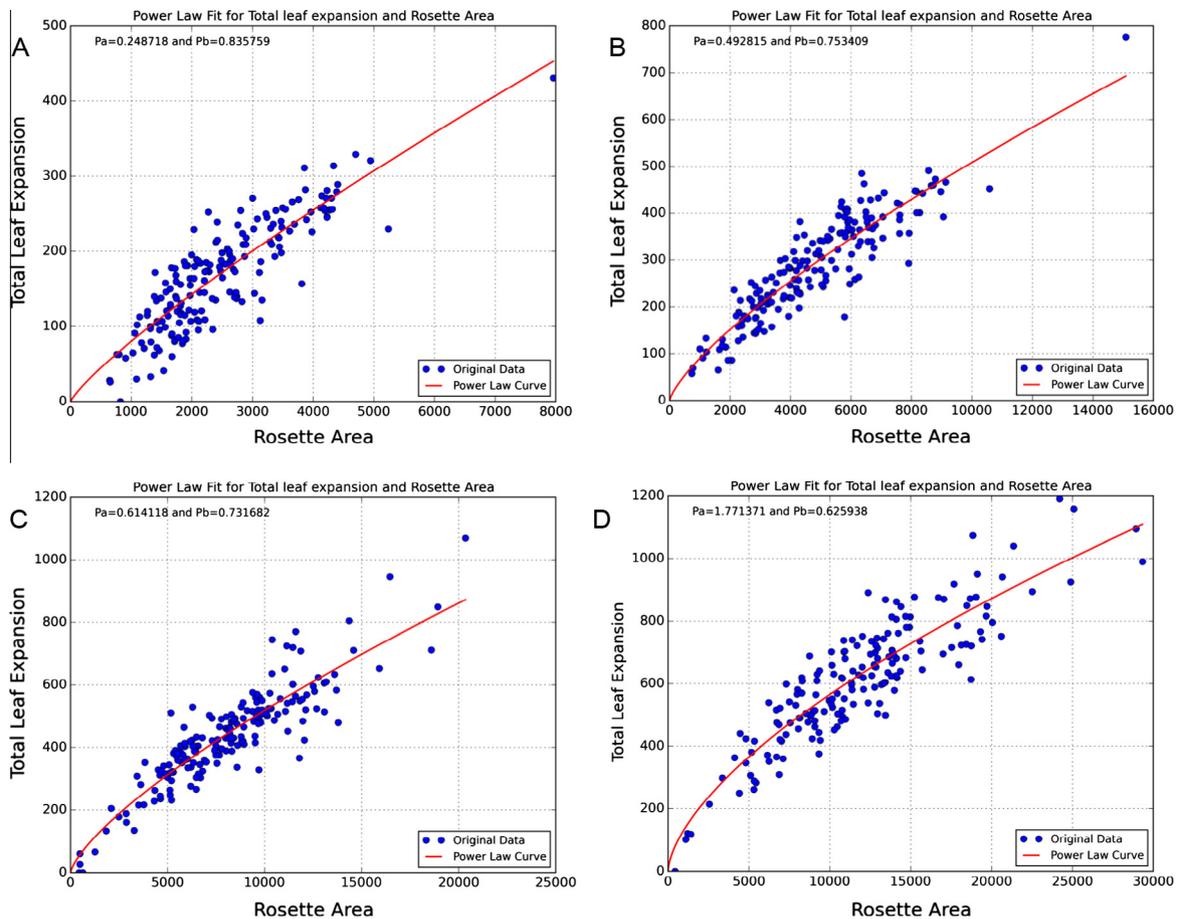


Fig. 15. Relationship of rosette area and total leaf expansion for indoor environment. The relationship between rosette area and total leaf expansion fitted with a power law function. The exponents of the power law function decrease from Day 1 (the eighth day after plant emergence) to Day 4 (the fourteenth day after plant emergence). At the early developmental stage, the exponent is close to 1 (a linear relationship), and at the late developmental stage, the exponent is close to 0.5, matching a previous study.

and efficient for finding the leaf tips when the leaves were damaged, as was the case in Fig. 14 example.

3.3. Relationship between rosette area and total leaf expansion

3.3.1. Relationship for indoor pipeline

The estimated power law exponent, b , from observation date 1–4 is 0.836 (Fig. 15A), 0.753 (Fig. 15B), 0.750 (Fig. 15C), and 0.636 (Fig. 15D), respectively, indicating a change in shape from near-linearity to a more curved relationship. The distribution of bootstrap estimates for a and b obtained from 10,000 simulations are shown in Fig. 16. The mean, median, and mode for a and b , respectively, are very close to the original estimate from the least squares fit.

Because the confidence limits for the four observation dates do not overlap, it can be said that b decreases with time while the slope, a , factor increases, although there is considerable uncertainty about the values of a . However, the range of values for b is within the reasonable expectations of being not greater than 1 and not less than 0.5. These values are consistent with the idea young leaves mainly grow by elongation. However, in later developmental stages, total leaf expansion slows relative to leaf-width growth increases, which become the main contributor to increasing rosette area. Chitwood et al. (2012) reported a linear relationship between total leaf length and square root of the total leaf area (i.e., $b = 0.5$) for tomato leaves at a late developmental stage. This

result is very similar to our finding of $b = 0.636$ during late development.

This dynamic relationship between total leaf expansion and rosette area has not been reported previously, quite possibly because destructive sampling made it impossible to collect time-series data from the same plant during growth. Our non-invasive imaging method, however, can be used to track the time-series development pattern for single plants in genetic mapping populations.

3.3.2. Relationship for field pipeline

The estimated power law parameters a and b and curves for observation dates 1–6 in the field are shown in Fig. 17A–F, respectively. The b values also followed a descending trend over time. Fig. 18 shows the analyses of the parameter distributions and the bootstrap 95% confidence intervals. This result follows what is seen in the chamber; that is, b is close to 1 at the beginning of the growth period but decreases over time. As above, the confidence intervals of the first and last days do not overlap, showing that a increases with time and b falls. Again, b showed the pattern of being close to 1 earlier and not less than 0.5 later. This also appears similar to the Chitwood et al. (2012) result.

It is noticeable that the b value increased on the last observation date and that there is an high-valued outlier (circled in Fig. 17F). Fig. 17G shows that this outlier was from a bolting plant that yielded a confusing single-rosette binary image. The center part of this binary was missing due to bolting, and a noisy plant outline

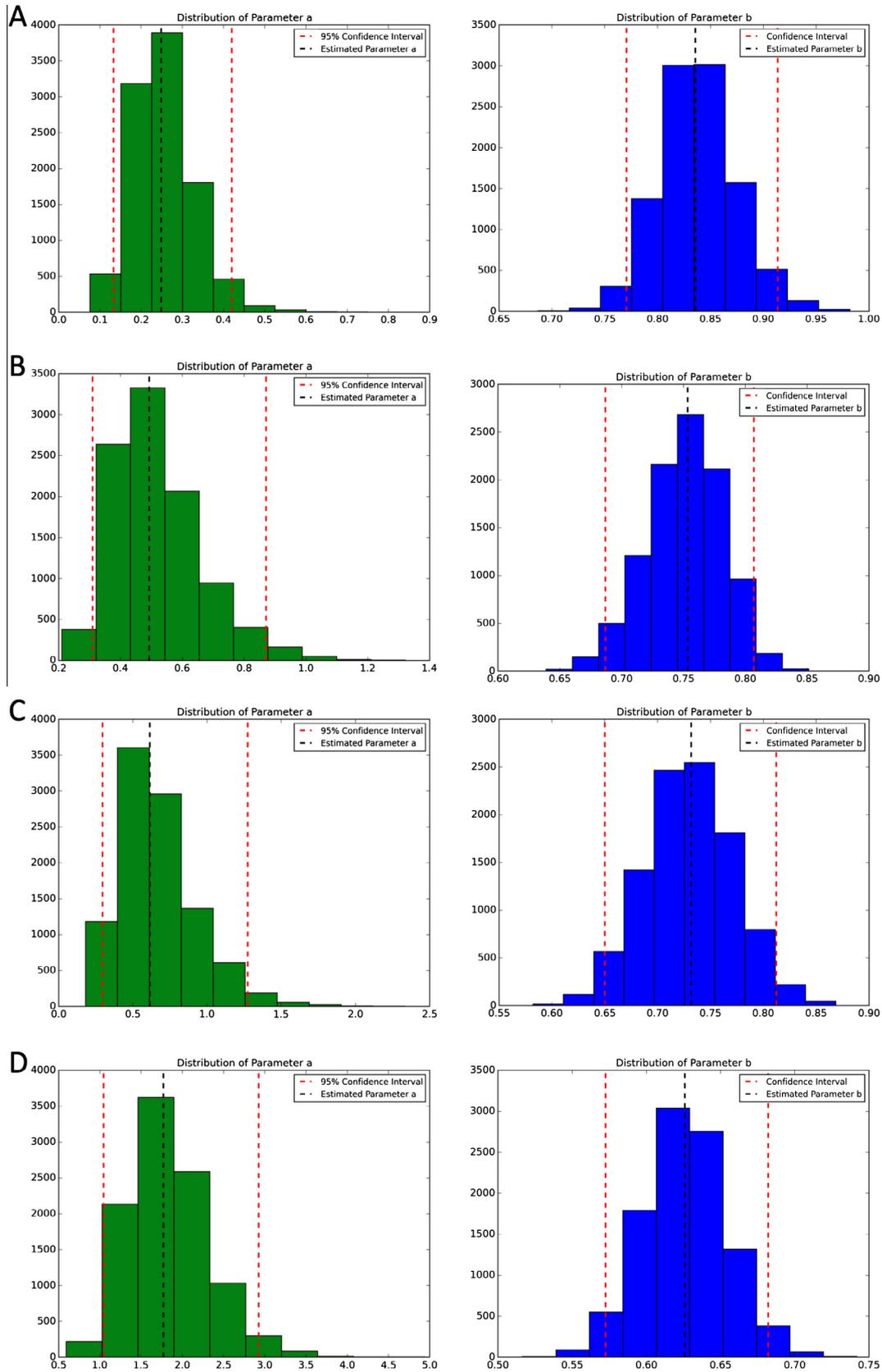


Fig. 16. Parameter estimations for indoor environment. The distribution histograms of the parameters *a* and *b* of four different time points. Figure (A)–(D) stands for Day 1 to Day 4, respectively. The green histograms are the distributions for parameter *a* and the blue histograms are for parameter *b*. Red-dashed lines stand for 95% confidence intervals and black-dashed lines are estimated parameters. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

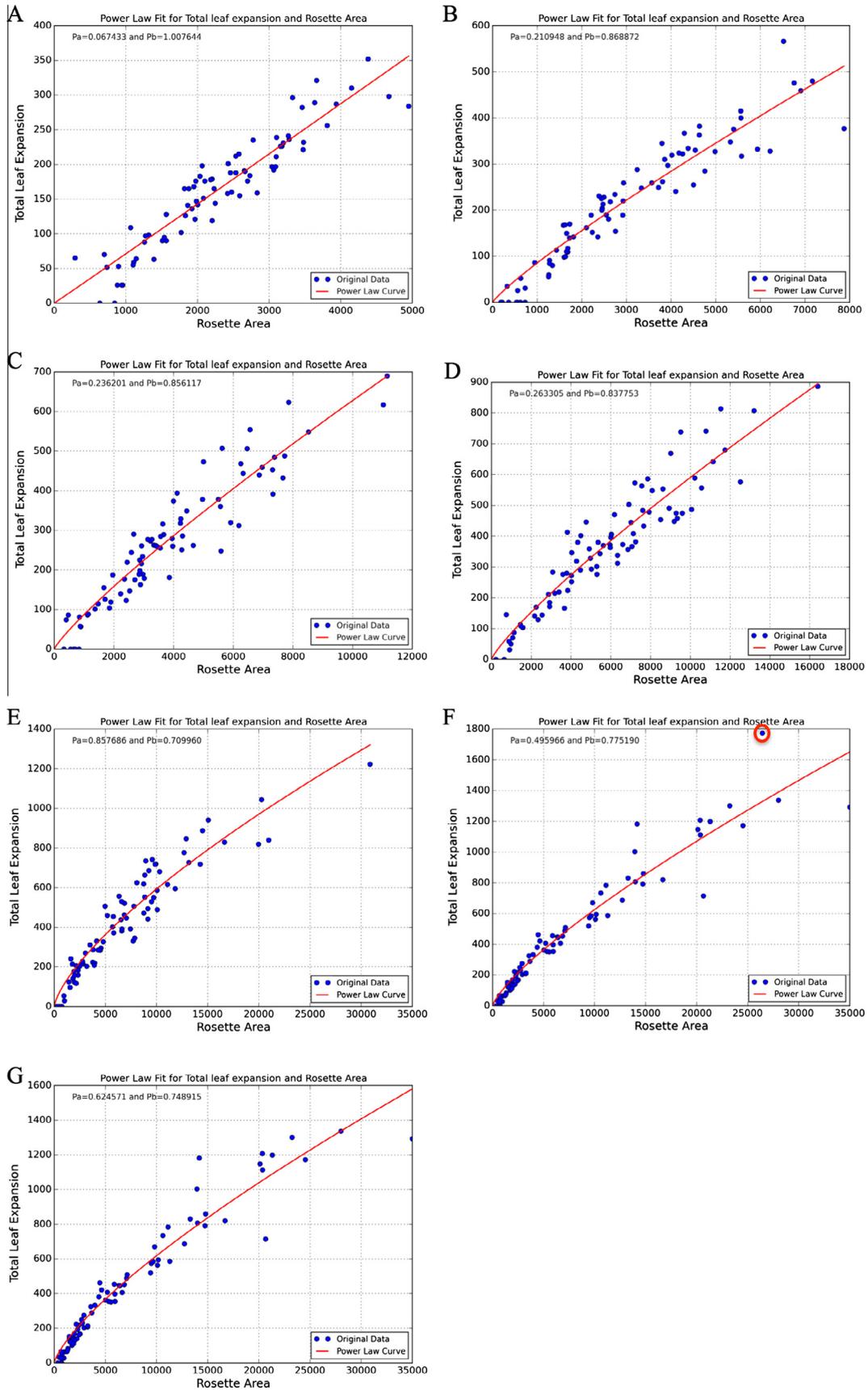


Fig. 17. Relationship of rosette area and total leaf expansion for field environment. The highlighted data point in (F) is an outlier. (G) is the same as last day after outlier removed.

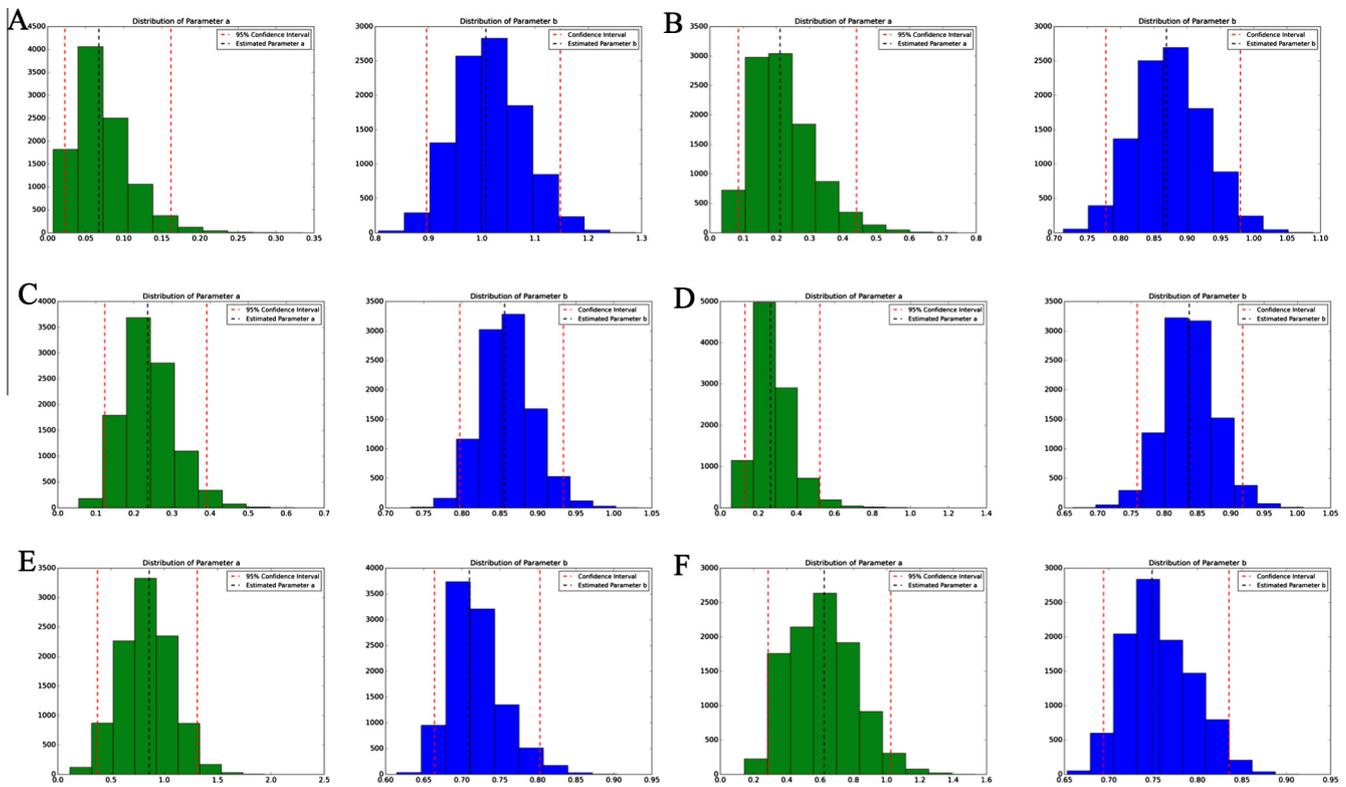


Fig. 18. Parameter estimations for field environment. The analyses of the parameter distributions and the bootstrap 95% confidence intervals.

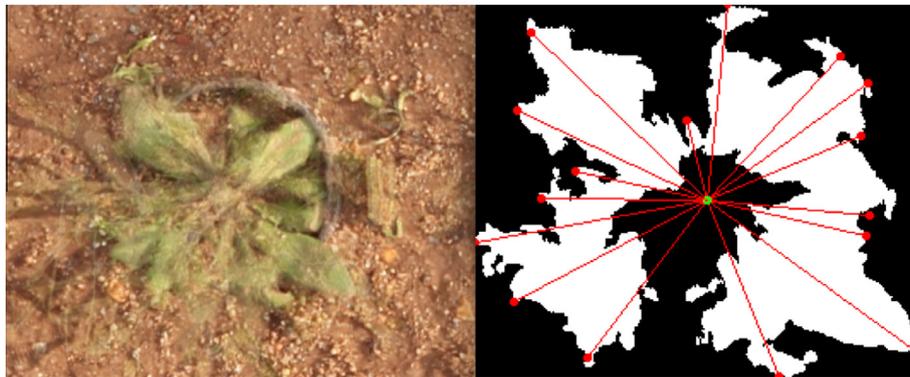


Fig. 19. Original image and analyzed image of the outlier. Highlights outlier from Fig. 17F. This plant started bolting on Day 6, which led to a confusing orthophoto. The segmented binary image missed the plant's central section and false leaf tips were detected, which caused reduced rosette area but increased total leaf expansion.

caused falsely-detected leaf tips (Fig. 19). After removing this outlier, the b value dropped to 0.749.

It is possible the parameter b would keep decreasing if a few more time points could be used. However, after the sixth observation date, many genotypes started bolting, which created difficulties in generating clean binary images and decreased the reliability of the total leaf expansion and rosette area measurements.

4. Conclusion and perspectives

Our HTP pipeline deals with many photogrammetric issues that have not been previously considered in most biology studies.

Images were first processed for optical and perspective distortion removal to make sure the true geometric quantities (e.g., distance and area) could be measured based on 2D images. Next, mosaicked orthophotos of each shelf were created hourly for a mapping population. Last, segmented single-rosette images were extracted from the orthophotos for measuring leaf length and rosette area with our novel image processing algorithm. This method transforms computing-expensive image processing operations to mathematical curve fitting problems and provides a reliable solution for leaf tip detection in the face of leaf irregularities, segmentation errors, and damage.

Our imaging system and pipeline is capable of being scaled up for larger indoor environments, such as greenhouse imaging systems. More plants can be imaged within each image frame without

losing any details by using newer digital cameras with wide-angle lens. However, the current pipeline throughput might be limited by the bottleneck of the data transfer rate between cameras, cluster nodes and data storage, but this can readily be minimized by adding faster network hardware and more cluster nodes.

With this framework of phenotyping, time-series phenotypic traits of a mapping population can be extracted and analyzed in a short period of time. In this study, we found a power-law correlation between total leaf expansion and rosette area from our time-series analysis. At the early developmental stage, this relationship was close to linear; toward the end of the developmental stages, the exponent started decreasing. The late-stage finding is very similar to the results reported by Chitwood et al. (2012).

There are still some aspects of the current pipeline that need to be improved for future work. The color-dot system of chamber study was used to track flat position and orientation changes in the orthophoto. However, the QR-code approach proved its robustness and feasibility for storing metadata and could be competitive in chamber settings to enable genotype assignment. Moreover, aerial-level field studies could also use QR-code system to store crop variety or field position metadata.

The facility with which cameras can be currently swapped out by changing a few interface scripts begs the question if this can be extended to thermal or multi-spectral imagers. An issue with thermal cameras is their low resolution. A key step in orthophoto generation is the detection of common points in different images using the SIFT algorithm (Lowe, 1999). This method exploits small, granular image features that are not typically present in low-resolution images. However, this problem can be expected to ameliorate as thermal camera resolutions improve so the pipeline reported here can be considered to be prepositioned to take advantage of such advances. The situation is different for multi-spectral cameras that image selective channels (more than RGB but generally not many) at high spatial resolution. To incorporate such cameras, white balancing would not be processed from the interface script; instead it would be done using reference panels during the camera spectral calibration step that precedes imaging. Lens distortion removal would remain unaltered and TIFF files readily handle multi-channel images. PhotoScan would need to process channels in groups of three but this would be compensated for by increasing the number of compute nodes as was suggested for larger greenhouse studies. Thus, multi-spectral cameras do not appear to present insuperable challenges.

To conclude, we have presented a low-cost, versatile, and high-throughput phenotyping pipeline based on imaging technology to quantify plant phenotypic traits under different environmental conditions (both laboratory and field). Except for a minor step in the field version, our pipeline integrates a series of fully automated operations, including: image data acquisition, image data transfer and management, genotype assignment, phenotypic trait extraction, and analytical processes. The use of interface scripts enabled different camera types to be readily exchanged. This flexibility can extend to multi-spectral cameras and, as resolution improves, to thermal cameras as well, thus moving 3D sensing beyond the visible spectrum.

Acknowledgement

This study was funded by the National Science Foundation – United States Grant IOS-0923752 to Cynthia Weinig, Stephen Welch, Julian Maloof, and C. Robertson McClung. We want to thank Smita Sharan (KSU), Rishab Vaswani (KSU), Lakshmi Narjala (KSU), Wen Fung Leong (KSU), Huan Wang (KSU), and Siben Li (University of Edinburgh) for providing valuable suggestions on modeling and statistics. We are also appreciative of the helpful comments and suggestions of reviewers.

References

- Ali, M.M., Al-Ani, A., Eamus, D., Tan, D.K., 2012. A new image-processing-based technique for measuring leaf dimensions. *Am-Eurasian J. Agric. Environ. Sci.* 12, 1588–1594.
- Arabidopsis Genome Initiative, 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408 (6814), 796–815.
- Arvidsson, S., Pérez-Rodríguez, P., Mueller-Roeber, B., 2011. A growth phenotyping pipeline for *Arabidopsis thaliana* integrating image analysis and rosette area modeling for robust quantification of genotype effects. *New Phytol.* 191 (3), 895–907.
- Bolon, Y.T., Haun, W.J., Xu, W.W., Grant, D., Stacey, M.G., Nelson, R.T., Gerhardt, D.J., Jeddeloh, J.A., Stacey, G., Muehlbauer, G.J., Orf, J.H., Naeve, S.L., Stupar, R.M., Vance, C.P., 2011. Phenotypic and genomic analyses of a fast neutron mutant population resource in soybean. *Plant Physiol.* 156 (1), 240–253.
- Bongaarts, J., 2014. United Nations, department of economic and social affairs, population division, sex differentials in childhood mortality. *Popul. Dev. Rev.* 40 (2), 380.
- Bylesjö, M., Segura, V., Soolanayakanahally, R.Y., Rae, A.M., Trygg, J., Gustafsson, P., Jansson, S., Street, N.R., 2008. LAMINA: a tool for rapid quantification of leaf size and shape parameters. *BMC Plant Biol.* 8 (1), 82.
- Candela, H., Martínez-Laborda, A., Micol, J.L., 1999. Venation pattern formation in *Arabidopsis thaliana* vegetative leaves. *Dev. Biol.* 205 (1), 205–216.
- Chen, D., Neumann, K., Friedel, S., Kilian, B., Chen, M., Altmann, T., Klukas, C., 2014. Dissecting the phenotypic components of crop plant growth and drought responses based on high-throughput image analysis. *Plant Cell* 26 (12), 4636–4655.
- Chitwood, D.H., Headland, L.R., Filiault, D.L., Kumar, R., Jiménez-Gómez, J.M., Schragger, A.V., Park, D.S., Peng, J., Sinha, N.R., Maloof, J.N., 2012. Native environment modulates leaf size and response to simulated foliar shade across wild tomato species. *PLoS ONE* 7 (1).
- Cookson, S.J., Chenu, K., Granier, C., 2007. Day length affects the dynamics of leaf expansion and cellular development in *Arabidopsis thaliana* partially through floral transition timing. *Ann. Bot.* 99 (4), 703–711.
- Dornbusch, T., Lorrain, S., Kuznetsov, D., Fortier, A., Liechti, R., Xenarios, I., Fankhauser, C., 2012. Measuring the diurnal pattern of leaf hyponasty and growth in *Arabidopsis* – a novel phenotyping approach using laser scanning. *Funct. Plant Biol.* 39 (11), 860–869.
- Dornbusch, T., Michaud, O., Xenarios, I., Fankhauser, C., 2014. Differentially phased leaf growth and movements in *Arabidopsis* depend on coordinated circadian and light regulation. *Plant Cell* 26 (10), 3911–3921.
- Duda, R.O., Hart, P.E., 1972. Use of the Hough transformation to detect lines and curves in pictures. *Commun. ACM* 15 (1), 11–15.
- Fiorani, F., Schurr, U., 2013. Future scenarios for plant phenotyping. *Annu. Rev. Plant Biol.* 64, 267–291.
- Furbank, R.T., von Caemmerer, S., Sheehy, J., Edwards, G., 2009. C4 rice: a challenge for plant phenomics. *Funct. Plant Biol.* 36 (11), 845–856.
- Furbank, R.T., Tester, M., 2011. Phenomics – technologies to relieve the phenotyping bottleneck. *Trends Plant Sci.* 16 (12), 635–644.
- Granier, C., Aguirrezabal, L., Chenu, K., Cookson, S.J., Dauzat, M., Hamard, P., Thioux, J., Rolland, G., Bouchier-Combaud, S., Lebaudy, A., Muller, B., Simonneau, T., Tardieu, F., 2006. PHENOPSIS, an automated platform for reproducible phenotyping of plant responses to soil water deficit in *Arabidopsis thaliana* permitted the identification of an accession with low sensitivity to soil water deficit. *New Phytol.* 169 (3), 623–635.
- Green, J.M., Appel, H., Rehrig, E.M., Harnsomburana, J., Chang, J.F., Balint-Kurti, P., Shyu, C.R., 2012. PhenoPhyte: a flexible affordable method to quantify 2D phenotypes from imagery. *Plant Meth.* 8 (1), 1–12.
- Hunt Jr., E.R., Cavigelli, M., Daughtry, C.S., Mcmurtrey III, J.E., Walthall, C.L., 2005. Evaluation of digital photography from model aircraft for remote sensing of crop biomass and nitrogen status. *Precision Agric.* 6 (4), 359–378.
- Jansen, M., Gilmer, F., Biskup, B., Nagel, K.A., Rascher, U., Fischbach, A., Briem, S., Dreissen, G., Tittmann, S., Braun, S., De Jaeger, I., Metzloff, M., Schurr, U., Scharr, H., Walter, A., 2009. Simultaneous phenotyping of leaf growth and chlorophyll fluorescence via GROWSCREEN FLUORO allows detection of stress tolerance in *Arabidopsis thaliana* and other rosette plants. *Funct. Plant Biol.* 36, 902–914.
- Lowe, D.G., 1999. Object recognition from local scale-invariant features. *Computer vision, 1999. The proceedings of the seventh IEEE international conference on*, vol. 2. IEEE, pp. 1150–1157.
- McMullen, M.D., Kresovich, S., Villeda, H.S., Bradbury, P., Li, H., Sun, Q., Flint-Garcia, S., Thornsberry, J., Acharya, C., Bottoms, C., 2009. Simultaneous phenotyping of leaf growth and chlorophyll fluorescence via GROWSCREEN FLUORO allows detection of stress tolerance in *Arabidopsis thaliana* and other rosette plants. *Funct. Plant Biol.* 36 (11), 902–914.
- O'Malley, R.C., Ecker, J.R., 2010. Linking genotype to phenotype using the *Arabidopsis* unimutant collection. *Plant J.* 61 (6), 928–940.
- Otsu, N., 1975. A threshold selection method from gray-level histograms. *Automatica* 11 (285–296), 23–27.
- Pérez-Pérez, J.M., Serrano-Cartagena, J., Micol, J.L., 2002. Genetic analysis of natural variations in the architecture of *Arabidopsis thaliana* vegetative leaves. *Genetics* 162 (2), 893–915.
- Phillips, R.L., 2010. Mobilizing science to break yield barriers. *Crop Sci.* 50, S-99–S-108.
- Reynolds, M., Foulkes, M.J., Slafer, G.A., Berry, P., Parry, M.A., Snape, J.W., Angus, W.J., 2009. Raising yield potential in wheat. *J. Exp. Bot.*, erp016

- Savitzky, A., Golay, M.J., 1964. Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem.* 36 (8), 1627–1639.
- Tchebychev, P.L., 1854. Théorie des mécanismes connus sous le nom de parallélogrammes. *Mémoires des Savants étrangers présentés à l'Académie de Saint-Petersbourg*, vol. 7, pp. 539–586.
- Tester, M., Langridge, P., 2010. Breeding technologies to increase crop production in a changing world. *Science* 327 (5967), 818–822.
- Walter, A., Scharr, H., Gilmer, F., Zierer, R., Nagel, K.A., Ernst, M., Wiese, A., Virnich, O., Christ, M.M., Uhlig, B., Jünger, S., Schurr, U., 2007. Dynamics of seedling growth acclimation towards altered light conditions can be quantified via GROWSCREEN: a setup and procedure designed for rapid optical phenotyping of different plant species. *New Phytol.* 174 (2), 447–455.
- Weigel, D., Mott, R., 2009. The 1001 genomes project for *Arabidopsis thaliana*. *Genome Biol.* 10 (5), 107.